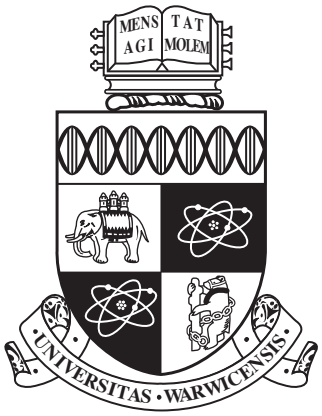


Computational models of morphology's effects on cellular dynamics

by

Faiz Sayyid

A thesis submitted to The University of Warwick

in partial fulfilment of the requirements

for admission to the degree of

Doctor of Philosophy

Department of Computer Science

The University of Warwick

September 2016

Contents

Acknowledgements

First and foremost, I would like to thank my supervisor Sara Kalvala for guiding me through my PhD and offering her constant enthusiasm for my work. I would also like to thank my colleagues, past and present, at the Department of Computer Science for helping me develop academically and professionally. Finally I wish to thank my family who have always encouraged me in all of my endeavours.

Sponsor

The research presented in this thesis was made possible by the support of:

- The University of Warwick, United Kingdom:
  Engineering and Physical Sciences Research Council Doctoral Training Partnership Studentship (2012–2016)

Publications

Parts of this thesis have been published by the author in the following:

Chapter 4

Faiz Sayyid and Sara Kalvala. 2016. On the importance of modelling the internal spatial dynamics of biological cells. *Journal of Biosystems*, 145:53 – 66. ISSN 0303-2647. doi: http://dx.doi.org/10.1016/j.biosystems.2016.05.012

Chapter 5

Faiz Sayyid and Sara Kalvala. 2016. On the importance of modelling the internal spatial dynamics of biological cells. *Journal of Biosystems*, 145:53 – 66. ISSN 0303-2647. doi: http://dx.doi.org/10.1016/j.biosystems.2016.05.012

Abbreviations

ATP Adenosine Triphosphate 31

BD Brownian Dynamics 78

BM Brownian Motion 47

BMR Bi-Molecular Reaction 261

CCD Charge-Coupled Device 53

CLE Chemical Langevin Equation 66

CME Chemical Master Equation 65

CNS Central Nervous System 13

CUDA Compute Unified Device Architecture 95

DNA Deoxyribonucleic acid 87

EMM Euler-Maruyama Method 70

ER Endoplasmic Reticulum 29

FCS Fluorescence Correlation Spectroscopy 52

FDM Finite Difference Methods 95

FRAP Fluorescence Recovery after Photobleaching 52

FSP Finite State Projection 69

GPGPU General Purpose Graphics Processing Unit Computing 255

IVP Initial Value Problem 73

LV Lotka-Volterra 84

LVD Lotka-Volterra-Diffusion 223

MD Molecular Dynamics 78

MP Measurement Point 134

NC Nuclear to Cytoplasm 224

ODE Ordinary Differential Equation 94

OP Original Point 134

PDE Partial Differential Equation 88

ReDi-Cell Reaction-Diffusion Cell 231

RK4 Runge-Kutta 4th Order Method 95

RKM Runge-Kutta Method 71

RRE Reaction Rate Equation 68

SC Structure to Cytoplasm 225

SDE Stochastic Differential Equation 70

SPT Single Particle Tracking 52

SSA Stochastic Simulation Algorithm 69

VCell Virtual Cell Simulation Package 106

VisIt Visualise It 88

Abstract

Spatial effects such as cell shape, internal cellular organisation and cellular plasticity have very often been considered negligible in models of cellular pathways, and many existing simulation infrastructures do not take such effects into consideration. However, recent experimental results and systems level theories suggest that even small variations in shape can make a large difference to the fate of the cell. This is particularly the case when considering eukaryotic cells, which have a complex physical structure and many subtle control mechanisms. Bacteria are also interesting for their variation in shape, both between species and in different states of adaptation.

In this thesis we perform simulations that quantify the effect of three aspects of morphology - external cellular shape, internal cellular organisation and processes that change the shape of the cell - on the behaviour of model cellular pathways. To perform these simulations we develop Reaction-Diffusion Cell (ReDi-Cell), a highly scalable General Purpose Graphics Processing Unit Computing (GPGPU) cell simulation infrastructure for the modelling of cellular pathways in spatially detailed environments. ReDi-Cell is validated against known-good simulations, prior to its use in new work.

By measuring reaction trajectories and concentration gradients we quantify the responses of simulated cellular pathways to these three spatial aspects. Our results show that model cell behaviour is the composite of cellular morphology and reaction system. Different reaction systems display different dynamics even when placed in identical environments.

Traditionally, computational approaches to cell biology have been focussed upon investigating how changes to reaction dynamics alter cellular behaviour. This thesis, on the other hand, demonstrates another way in which reaction dynamics can be altered, by changing the morphology of the cell.

CHAPTER 1

Introduction

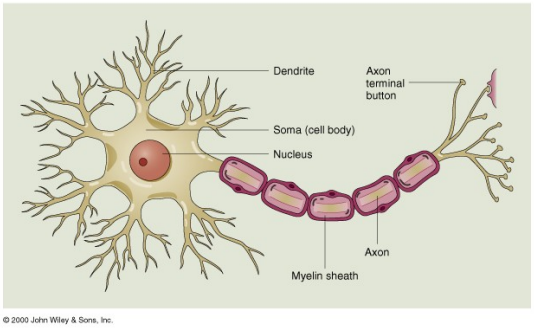
The simulation of living cells is important for both revealing the complex mechanisms behind their behaviour and predicting their future actions. *In silico* studies of the cell allow a degree of measurement and control that is impossible *in vivo* or *in vitro*. As the fundamental unit of life, the simulation of the behaviour of individual cells has been described as one of the grand challenges of the 21st century [Tomita, 2001]. Despite this, spatial effects such as external cell shape, internal organisation and processes that change the shape of the cell have very often been considered negligible in models of cellular behaviour, and many existing simulation studies do not take such effects into consideration. Recent experimental results are reversing this judgement by showing that spatial variations can make a big difference in the fate of a cell [Cowan et al., 2012; Ihekwaba et al., 2004; Kholodenko, 2009; Maeder et al., 2007; Mazel et al., 2009; Neves and Iyengar, 2009; Neves et al., 2008; Ohshima et al., 2012; Pham and Ichikawa, 2013; Terry, 2014; Terry and Chaplain, 2011; Wang et al., 2006; Zhao et al., 2011]. By combining data on the cell's chemical environment with morphological information, integrative modelling approaches allow us to explore the role of cell shape in cell behaviour. Although there exist spatial simulations of specific cells and specific systems within those cells [Means et al., 2006; Orton et al., 2005; Zhao et al., 2011], there has not yet been a concerted effort to systematically explore the cell's behaviour in response to its spatial configuration. Further work is required to understand the influence of spatial organisation on cellular processes [Kholodenko, 2009] - this thesis is concerned with exploring the nature of this influence.

1.1 Cell morphology

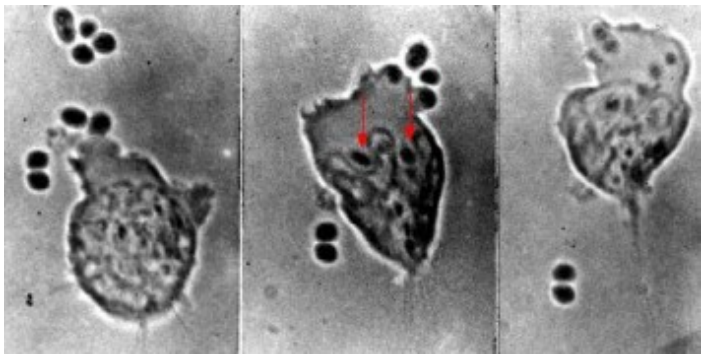
Cell morphology, that is the external shape of the cell coupled with its internal organisation, varies wildly throughout nature. The morphology of the cell is often vital to its function, with a range of literature attempting to connect the two [Shah, 2010]. For example, the regular building-block nature of epithelial formations hint at some sort of barrier function. Figure 1.1 shows some particularly famous examples of cells where shape is vital for function.

Figure 1.1(a) shows a diagram of a human neuron, the cell responsible for communication between the Central Nervous System (CNS), and the sense organs and muscles. The axon of the neuron is large compared to the rest of the cell, up to 12 m in the case of the giant squid. The neuron's morphology allows for high levels of connectivity to other nerve cells. The axon divides at the end to duplicate input to other neurons and the dendrites are capable of branching to such an extent as to allow for 100,000 inputs to a single neuron. Figure 1.1(b) depicts an immune cells known as a neutrophil. This figure shows the neutrophil changing shape to engulf and destroy several bacteria. Their shape shifting ability also permits them to move through the tightly packed cells of the host's tissues to reach their targets. Figure 1.1(c) shows a red mammalian blood cell, responsible for absorbing oxygen in the lungs and distributing it to other parts of the body. Their most noticeable morphological adaptation is their biconcave shape, which increases their surface area, allowing more oxygen and carbon dioxide to be exchanged per unit time. A perfectly flat cell would have an even better surface area to volume ratio. As this is not the case it is likely that the cell's shape is compromise between functions.

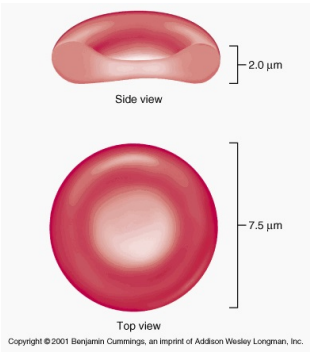
Living cells also take on a bewildering range of sizes; from the diminutive *Pelagibacter ubique* enclosing a volume of 0.01 μm^3 [Young, 2006] to the gargantuan *Valonia ventricosa*. Figure 1.2 shows a range of cells compared in size and shape to *E. coli*. The driving factors for cells adopting a particular size, from both an evolutionary and mechanistic point of view are still unclear. But



(a) Diagram of a neuron. Reproduced from anatomypicture.us Original credit unknown, published by John Wiley and Sons. Inc



(b) Photographs of a neutrophil engulfing pathogens. In the left pane a neutrophil extends part of its body around two *pneumococci*. In the centre pane the bacteria are engulfed, the red arrows show their location within the neutrophil. In the right pane two *pneumococci* have managed to escape from the neutrophil. Reproduced from theimmunology.com



(c) Diagram of a red blood cell. Reproduced from theartofmed.wordpress.com Original credit: Benjamin Cummings

Figure 1.1: Adaptive morphology of several cells, a neuron, a neutrophil engulfing a pathogen and a red blood cell

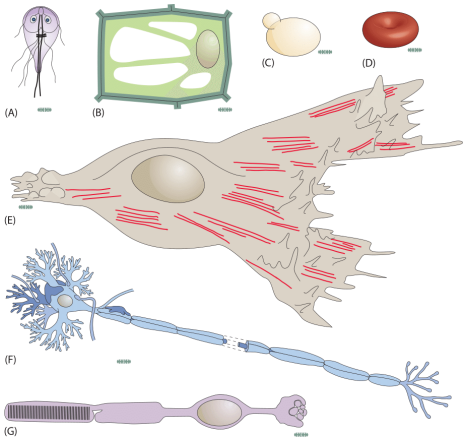


Figure 1.2: Cartoons of several different types of cells as compared to an *E. Coli* bacterium of length 1 /*mum*. (A) The protist *Giardia lamblia*, (B) a plant cell, (C) a budding yeast cell, (D) a red blood cell, (E) a fibroblast cell, (F) a eukaryotic nerve cell, and (G) a rod cell from the retina. Reproduced from [Milo and Philips, 2016]

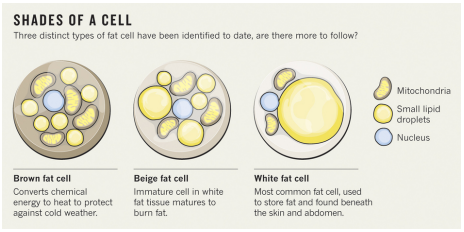


Figure 1.3: Differences in substructure distribution in three different types of fat cells. Reproduced from [Owens, 2014]

it is typically argued that greater surface to volumes ratios increase the rate at which nutrients can be delivered to the internals of the cell, whilst at the same time limiting the size of cell that has to be supported [Young, 2006].

Internal organisation can also differ markedly between cells. Some cells are incredibly simple, lacking any real internal organisation, their insides form a "soup" of components whereas others are highly organised. Within organised cells there is a great variety of structure. Figure 1.3 shows two different types of fat cells; a white fat cell, a beige fat cell and a brown fat cell. Internal space is configured differently in these three cells, volume is concentrated in a few substructures in the white fat cell but highly distributed in the brown fat cell.

1.2 Cellular Dynamics

A cell's dynamics can be thought of as the totality of all spatially localised and temporally dynamic biochemical reactions that occur within the cell [celldynamics.johnshopkins.edu, 2014]. Many reactions may take place in a cell every second and control over these is necessary. In cells this is often accomplished with the use of *enzymes*, biological catalysts that speed up reactions without themselves being consumed [Alberts et al., 2013]. The majority of the reactions within the cell do not take place in isolation, but instead a number of reactions will form a step by step chain known as a biological pathway, with many pathways composing a biochemical network. Chemical pathways control the behaviour of the cell. For example a combination of chemical pathways (forming a network) might control the assembly of a fat or a protein, turning a gene on or off, or causing the cell to move. In order to simulate some component of cell behaviour we do not need a complete picture of cellular dynamics, that is, to know about *everything* that is taking place in the cell. Rather, we separate out the section that we intend to study, and talk about the cellular dynamics of a particular process or network.

In this thesis we simulate abstract reaction pathways that have the properties of *in vivo* pathways. We select three reaction systems that are mathematically representative of biological pathways found in nature; the Lotka-Volterra (LV) reaction system has been used as an abstraction of biochemical oscillations [Andrews et al., 2009], the Bi-Molecular Reaction (BMR) system is equivalent to a generic second order reaction system, and the Turing system has been used to model morphogenesis [Turing, 1990].

The reactions that constitute chemical pathways cannot take place if the chemicals that constitute them are not in the correct location. They must be transported to the locations within the cell where they are required - this accounts for the spatially localised facet of the definition of cellular dynamics. The mechanism used for transport within the cell is often diffusion, the processes by

which matter is moved from one part of a system to another down a concentration gradient[Crank, 1979]. Any process that requires the shuttling of chemicals across the cell will most likely utilise diffusion, which requires less energy than other transport mechanisms, but is slower than other cellular transport mechanisms for larger molecules [Alberts et al., 2013]. Reactive and diffusive processes act in concert which results in cellular dynamics - the behaviour of the cell.

1.3 Cell dynamics and cell morphology

Computer simulation and traditional lab experimentation are making it increasingly clear that the reaction pathways and diffusive processes that constitute cellular dynamics are affected by the morphology of the cell [Kholodenko, 2009]. By adding space and hence inviting diffusion into reacting biochemical systems, *gradients* of concentration across space, rather than constant values in one position are introduced. The varying concentration induced by the gradient causes changes in the outcome of biochemical pathways. Concentration gradient formation is an important field of study in cell biology and it has been established that the size, the shape and internal organisation of the cell all play a pivotal role in inducing the concentration gradient and thus affect pathway behaviour [Meyers et al., 2006; Neves et al., 2008; Pham and Ichikawa, 2013]. Concentration gradients have been recorded in a variety of pathways including NF-kB, MAPK, GAP and mitosis [Fuller et al., 2008; Maeder et al., 2007; Niethammer et al., 2004] and there is growing evidence that some biological pathways could not take place without their existence. The neuron is a particularly striking example of cell that relies heavily on concentration gradients for its operation. The action potential in the axon could not take place without a very strong gradient of charge.

It is obvious that changing the size of the reaction vessel, in this case the cell, would induce different concentration gradients; the further a chemical has to travel the more likely it becomes that it will have undergone some sort of

change on its journey [Meyers et al., 2006]. Further evidence of the sensitivity of cell dynamics to cell size is found in numerical simulations of the yeast cell - cellular dynamics are markedly different in virtual yeast cells of different size [Zhao et al., 2011]. Experimental results have shown that changing the shape of the cell can also affect the formation of concentration gradients [Cowan et al., 2012; Neves and Iyengar, 2009; Neves et al., 2008].

Internal organisation and positioning have also been implicated in the generation of concentration gradients. This can occur by two mechanisms; **(i)** chemical components can be spatially restricted within the cell, **(ii)** the presence of a substructure could alter the formation of an existing concentration gradient.

1.4 Reaction-Diffusion models capture cell dynamics

When simulating the dynamics of living cells we seek to model the processes of reaction and diffusion undergone by chemicals within the cell. When models for both of these processes are combined we achieve the reaction-diffusion model, that is able to predict species concentrations in time and space.

As discussed in the background chapters 2 and 3, reaction-diffusion models have enjoyed considerable success in the modelling of cellular dynamics [Meyers et al., 2006; Soh et al., 2010; Terry and Chaplain, 2011]. This is due to their intrinsically spatial nature - they are well suited to modelling the arbitrarily shaped boundaries of in silico representations of cells. If a reaction-diffusion model is capable of describing *in vivo* behaviour then, by varying model parameters, it is possible to computationally identify the physical quantities of the cell and involved chemicals. This is of course contingent on a reaction-diffusion model being the correct choice.

Reaction-diffusion models have also been applied to morphogenesis, the biological process that causes an organism to take on a particular structure

[Youngson, 2006]. This was demonstrated in Alan Turing's seminal 1952 work, "The Chemical Basis of Morphogenesis" [Turing, 1990]. Patterns generated by the Turing reaction-diffusion system (along with others) have been found to mimic naturally occurring patterns of pigmentation and structure [Turing, 1990]. Reaction-diffusion models are also applied outside of biology, in fields as varied as geology [Bischoff and Piper, 2013] and crime hotspot prediction. [Short et al., 2010].

It is important to distinguish between reaction-diffusion *models* and reaction-diffusion *systems*. Whilst a reaction-diffusion model refers to any mathematics that attempts to capture reaction and diffusion, a reaction-diffusion system refers to particular, parametrised set of reacting and diffusing chemicals, that is a *specific* system.

Reaction-diffusion models can be formulated and solved in a variety of ways - stochastic simulations, differential equations and particle based methods have all been successfully applied. In this thesis we take a differential equation approach, using Ordinary Differential Equation (ODE)s for the modelling of reactions and Partial Differential Equation (PDE) for modelling diffusion. Chapter 3 discusses the mathematical formulation and solution of a variety of reaction-diffusion models.

1.5 Problem statement and contributions

In this thesis we systematically explore the effect that spatial configuration has on virtual biological pathways. Before we can do this we need to define what is meant by spatial configuration. In this thesis the three aspects of spatial of configuration explored are:

- The shape of the cell
- Internal cellular organisation
- Processes that change the shape of the cell

Specifically we state the research question as: **In what way does the configuration of cellular morphology alter the dynamics of numerically simulated biological pathways?**. To answer this overall research question we break it into three smaller research questions, answering them in chapters 5, 6 and 7.

- Research question 1:

  In what way does the surface shape of the cell alter the behaviour of simulated biological pathways?

- Research question 2:

  In what way does the internal organisation of the cell alter the behaviour of simulated biological pathways?

- Research question 3:

  In what way do processes that changes the shape of the cell alter the behaviour of simulated biological pathways?

In answering the research question, this thesis makes the following contributions:

- **Quantifying the responses of simulated biological pathways to common bacterial cell shapes**

  Using a numerical simulation package developed for this thesis and detailed in Chapter 4, we simulate virtual pathways in three common bacterial shapes. Bacterial cells take on a bewildering variety of external shapes [Young, 2006], but the three most common are *cocci* - spherically shaped, *bacilli* - rod shaped, and *spirilli* - helically shaped. We use our numerical simulation package to conduct novel simulations that demonstrate the effect that these common bacterial shapes have on the behaviour of model cellular pathways. Pathway wavefront shape, pathway concentration gradients, and chemical species distribution are measured in the

three different shapes. Through these results we show that variations in the shape or configuration of these common cell shapes alter model cell dynamics.

- **Quantifying the responses of simulated biological pathways to common configurations of internal cellular organisation**
  Using numerical simulations we examine the response of virtual pathways to internal cellular organisation. The nisinternal orgaation of cells can differ between different cell types. We use novel numerical simulations to explore the effect that three aspects of internal cellular organisation have upon model pathway dynamics. The three aspects investigated are: Nuclear to Cytoplasm (NC)ratio, nuclear position and spatial distribution of internal cellular volume. Species concentration, concentration gradients, wavefront shapes and translocation (travel) times are measured to quantify the sensitivity of pathway behaviour to internal organisation. Through these results we show how these different aspects of internal organisation alter model pathway dynamics.

- **Quantifying the responses of simulated biological pathways to common processes that alter morphological configuration**
  Our final simulations explore how processes that change the shape of simulated cells alter model pathway behaviour. External cell shape and internal organisation rarely remain the same through the cell's life cycle. Cells change their volume as they grow and their shape as they divide, but there are a number of other processes by which cells change their morphology. Non volume conservative processes include hypertrophy - an increase in size and atrophy - a decrease in size. Volume conservative processes can include changing shape for movement (in the case of a cell adhered to a surface) or elongation, in response to nutritional stress or to move through other cells. We use novel numerical simulations to show how changing the shape of common bacterial cells and amoebal cells alters virtual pathway

dynamics.

1.6 Structure of this thesis

The rest of this thesis is structured as follows:

Chapter 2 provides the necessary background on cell biology and its simulation together with a brief history of the role of computation. The classification and composition of living cells, cellular dynamics, and reaction-diffusion models of cellular dynamics are discussed in this section. These topics are discussed in the context of previous work within the field of cell simulation.

Chapter 3 presents information on the different techniques for formulating and solving reaction-diffusion simulations. The chapter begins with a discussion of the reaction-diffusion master equation and continues by both detailing different approximations to this formulation of a system and solution methods. Stochastic simulation, ODEs, PDEs, Stochastic Differential Equation (SDE)s and particle based simulations are all discussed. We pay particular attention to the strategies used in this thesis for simulation, namely PDE-ODE models.

Chapter 4 details the simulation, Reaction-Diffusion Cell (ReDi-Cell), used for producing the results of this thesis. We begin this chapter with a review of other cell simulation packages exploring their simulation capabilities. This review identifies a gap within the set of simulations; there is no open source simulation environment capable of scaling in performance to take advantage of end-user hardware. To this end we developed ReDi-Cell, a highly scalable GPGPU cell simulation infrastructure for the modelling of cellular pathways in spatially detailed environments. It is the first General Purpose Graphics Processing Unit Computing (GPGPU) PDE-ODE cell simulation infrastructure specifically designed for investigating the effect

of realistic cell shapes and internal components. The chapter continues with a demonstration of the capabilities of ReDi-Cell. We show ReDi-Cell's ability to approximate living cell shapes in chapter 4, justify the numerical methods chosen and explore its visualisation capabilities. Next we detail and justify the use of two reaction systems, the LV and the BMR system, that are used as abstractions for cellular reactions in the simulations within this thesis. The chapter ends with the validation of ReDi-Cell by comparing it against a peer review single cell simulation, Virtual Cell Simulation Package (VCell) [Loew and Schaff, 2001]. Results from VCell have been used in many publications.

Chapter 5 answers research question 1. In this chapter we investigate the role of external cell shape upon cellular dynamics. Simulations are performed to measure the effect of three common bacterial shapes (*Cocci*, *Bacilli* and *Spirilli*) on the behaviour of the LV and BMR model cellular pathways. Finally we demonstrate the sensitivity of a Turing system to abstract obstructions within the virtual cell. Pathway wavefront shape, pathway concentration gradients, and chemical species distribution are measured in the three different model cell shapes. The results in this chapter show how variations in external shape between these common cell shapes alter model cell behaviour.

Chapter 6 answers research question 2. This chapter is concerned with modelling the effect of internal cell organisation upon cellular dynamics. The literature establishes the existence of a link between the internal organisation of a living cell and an effect on a functional property, yet it does not characterise the relationships between the different aspects of internal organisation and their individual behavioural effects. We perform simulations that investigate the dynamical effects upon the model LV and BMR pathways, of three common metrics of internal cellular organisation: NC, nuclear position and substructure distribution. The values that

these metrics (NC, nuclear positioning and substructure distributions) are permitted to take are informed by real cells. In these simulations pathway wavefront shape, pathway concentration gradients and reaction trajectories are recorded. By performing simulations that systematically explore the effect of common metrics of internal organisation, we sharpen our view of how intracellular structure impacts pathway regulation in a general sense.

Chapter 7 answers research question 3. These results answer the question of how processes that change the shape of the cell affect cell dynamics. In this chapter we model two different types of morphological change: those in which volume is conserved and those in which volume is *not* conserved. A model *coccus*, *bacillus*, *spirillum* and irregularly shaped cell are permitted to increase their size, modelling cell growth and hypertrophy. A model *coccus* at the point of division is also simulated. Processes in which cells change their shape but conserve their volume are then simulated. Simulations measure concentration time series, heatmaps of concentration and wavefront shapes.

Chapter 8 concludes the thesis by summarising its contributions and drawing general conclusion from the chapters 4, 5 and 6. Further work in the area is also suggested.

CHAPTER 2

Cellular Structure and Behaviour

The cell is the fundamental unit of life capable of existing and replicating independently [Alberts et al., 1995]. All known life [1] is composed of biological cells, which provide structure and carry out the basic processes of life. There are a great many different varieties of cell, each with form and function specialised to their role, but all share some general properties [Alberts et al., 2013]. All cells consist of a membrane which encloses a gel-like substance known as the cytoplasm. The cytoplasm may in its turn contain cellular machinery. This cellular machinery contains entities such as biomolecules, nucleic acids, organelles and other structures. Organelles can be thought of as small machines which specialise in specific activities necessary for cell life.

2.1 The development of computational and quantitative approaches to cell biology

Naturally occurring cells were eventually classified into two types: Prokaryotes, where Deoxyribonucleic acid (DNA) is not partitioned from the cytoplasm, and Eukaryotes, where DNA is separated from the cytoplasm [Martin, 1977; Nelson et al., 2008]. Eukaryotes exist in both unicellular (such as amoeba) and multicellular organisms (such as plants and animals), where the total quantity of cells may be as high as 10^{13} for an organism such as a human [Alberts et al., 2013].

The first observation of the cell came in 1665 from Robert Hooke who named his discovery after the living accommodation of Christian Monks. Shortly after

[1] with the exception of the virus

this Antonie van Leeuwenhoek observes a living cell with moving components [Bisceglia, 2016]. It was not until 1885 that a theory of cells and their nature was described by Matthias Jakob Schleiden and Theodor Schwann [Karp, 1979]. The theory states that:

- All living organisms are composed of one or more cells
- The cell is the basic unit of structure and organization in organisms.
- Cells can only arise by division from pre-existing cells

Until the mid 1960s the majority of advances in cell biology were achieved through more and more powerful microscopes. After insulin was sequenced by Frederick Sanger it became clear that comparing sequences of DNA would become infeasible. It took nearly 15 years to develop the databases and algorithms required to store and analyse this sort of sequence information. In time this new field would become known as *Bioinformatics*. This new interest in using computational power to analyse sequence information led biologists to apply it to other areas of cell research. As well as asking "What does this data mean?" cell biologists began to ask "What is the mechanism behind cell behaviour?".

An explosion in computational power and new mathematics allowed previously formulated, and previously intractable mathematical models of cellular systems to be solved or approximated through simulation on digital computers. This increase in computational power coincided with the application of the newly developing field of complex systems to biology. Briefly stated, complexity theory states that rich, collective dynamic behaviour at the macrscopic level can result from simple interactions at the microscopic level and, conversely, that complex interactions at the microscopic level can give rise to simple behaviour at the macroscopic level [Rickles et al., 2007]. Computer programs that take a complex view of the cell allow the integration of many different simulations of different cellular processes to simulate a total emergent behaviour. Ever more complex integrative models of cells have been developed [Kholodenko, 2006] with

a complete model of the bacterium *M. Genitalium* predicting physical attributes from genetic information [Karr et al., 2012].

Still, the study of the cell is not yet complete, but integrative modelling approaches to cell biology are thought to hold the key to understanding cell behaviour in greater detail. Future advances in the field will require more data on the chemical and physical organisation of the cell from ever more advanced imaging equipment. Complete modelling of the cell is described as a grand challenge of the 21st century [Tomita, 2001].

An integrative simulation based approach to cell biology allows us to ask if our theories of cellular behaviour are correct. In this thesis we attempt to refine our theories of how cellular organisation alters behaviour by integrating cellular processes and spatial information. This work would be impossible *in vivo* or *in vitro* as we cannot exert a fine enough control over cell parameters.

2.2 Unifying Cellular Structure

As mentioned previously all cells share some common properties; they are all enclosed by a plasma membrane which encloses the cytoplasm, which in its turn, contains the cell's internal components.

2.2.1 The Cell Membrane

The cell's membrane is commonly referred to as the plasma membrane. It encloses all other components of the cell and acts as the gatekeeper to the inside of the cell, selectively allowing biomolecules to pass through in both directions. The plasma membrane allows the cell to concentrate nutrients absorbed from the environment and products of internal synthesis so that they can be used in the chemical reactions the cell performs. Without the plasma membrane the cell could not maintain its highly organised chemical environment necessary for supporting life [Alberts et al., 1995].

The 5 nm thick cell membrane is studded with selective channels and pumps

that give the membrane its partial permeability, allowing nutrients to enter, waste products to leave and reaction products to be contained. The membrane also plays host to proteins that allow signals to be transmitted and received by the cell. Despite its delicate structure the cell membrane is quite robust, resealing itself after rupture, plastically deforming, and growing with the cell. The number of membranes varies between cells - some simple cells posses only the exterior membrane, but complex Eukaryotes can contain many internal membranes that form the external barriers of the cell's organelles.

The lipid (fat-like) molecules that form the cell membrane have the remarkable property of being *amphiphillic*; each molecule consists of two parts, one of which is hydrophobic (water-insoluble) and the other hydrophillic (water-soluble). When placed in water these molecules self-organise to place as much of their hydrophobic portions in contact with each other as is possible, and simultaneously, as much of their hydrophillic portions as is possible in contact with the water. The specific amphillic molecules that make up the cell membrane are shaped such that when exposed to water they form a bilayer, creating small enclosed compartments called vesicles. As a general rule the hydrophobic tails of the membrane molecules are long chain hydrocarbons of the form $-CH_2 - CH_2 - CH_2-$. The way in which these amphiphillic molecules self-organise when exposed to water is an example of complex behaviour.

2.2.2 The Cytoplasm

The cytoplasm is the water based environment that is contained by the cell membrane. It consists of two parts: the cytosol and organelles and other structures [O'Connor et al., 2010]. The cytosol contains so many molecules packed together so tightly that, mechanically, it behaves like a gel [Fels et al., 2009]. The cytosol is also the site of a large portion of the cell's requisite chemical reactions. In both Prokaryotes and Eukaryotes the cytosol has been identified as the location for a large proportion of metabolic processes [Hoppert and Mayer, 1999]. Protein synthesis also takes place in the cytosol and in mammals the

majority of cellular proteins are localised within the cytosol [Foster et al., 2006]. The cytosol permits both reaction and diffusion of chemicals and as such is also responsible for transduction of externally received signals to the organelles or the nucleus.

2.2.3 Ribosomes

Ribosomes are the sites of protein synthesis in the cell. Responsible for the "grainy" effect in electron microscopy images of the cell, they commonly number in the thousands for common cells with more ribosomes present for cells with higher growth rates. In eukaryotes ribosomes may be attached to the rough Endoplasmic Reticulum (ER), but in prokaryotes they are distributed throughout the cytosol. Ribosomes are involved in the synthysis of proteins [O'Connor et al., 2010].

2.2.4 Cytoplasmic Inclusions

Cytoplasmic inclusions are chemical aggregations found in both prokaryotes and eukaryotes [Madigan et al., 2010]. Common examples of inclusions are glycogen particles in brown fat cells and muscle cells, lipid droplets in both brown and white fat cells. An exotic example of this structure are the gas vesicles, enclosed by a protein shell, found for example in marine cyanobacteria which are used to help position the cell within the water column. Another interesting inclusion is that of magnetic dipole displaying magnetosomes - particles of magnetite Fe_2O_3 enclosed by a membrane [Madigan et al., 2010]. These structures allow the bacteria to respond to the earth's magnetic field, and so it is hypothesised, direct the aquatic bacteria in which they are commonly found towards the ocean floor.

2.3 Eukaryotes

Biological cells exist in two forms: the nucleus containing eukaryote and the non nucleus containing prokayote [Nelson et al., 2008; Youngson, 2006]. In this section we discuss the eukaryote and nature of the organelle.

2.3.1 Organelles

Eukaryotes contain membrane bound organelles which serve to organise cellular functions. They are membrane bound intracellular compartments in which specialised activities, such as liberating energy for the cell, or processing chemicals during import or export. Despite the prokaryote's greater abundance in nature, eukaryotes are generally larger and display greater complexity [Alberts et al., 2013].

2.3.2 The nucleus

The nucleus is the defining feature of the eukaryote. Separated from the cytoplasm by a nuclear envelope consiting of two lipid bi-layers, it contains the hereditary material of the organism [Mai et al., 2005] in the form of DNA. DNA takes the form of a long polymer that encodes the instructions to build additional cellular matter inside the cell. In most eurkaryotes only a single nucleus is present, but in a few varieties of cell such as the mammalian red blood cells no nuclei are present. Cells with no nuclei can also result from cell division. Some types of cell including some types of protozoa [Zettler et al., 1997] and fungi [Horton, 2006] have more than one nucleus; these are examples of multinucleated cells.

The nuclear envelope is the interface between the nucleus and rest of cell. Like the external plasma membrane it is also selectively permeable and as such is able to regulate the flow of material between the nucleus and the cytoplasm. The membrane achieves this effect through holes in the envelope called nuclear pores. These pores allow RNA molecules through the envelope that carry instructions

from the DNA to the machinery for protein synthesis in the cytoplasm. By separating the instructions from the manufacturing, eukaryotes achieve a more precise control over themselves. The nuclear envelope also serves to separate material from the rest of the cell which is sometimes necessary in regulating a reaction inside the cytoplasm. In this case a participant of the reaction is translocated (transported) to the nucleus in order to reduce the speed of the reaction.

2.3.3 The mitochondrion and the chloroplast

Aside from the nucleus, two other organelles known as the mitochondrion and the chloroplast have important functions in the cell. Mitochondria are present in nearly all eukaryotic cells. They are known as the "power house" of the cell [Siekevitz, 1957] due to their role in providing energy for the cell in the form of Adenosine Triphosphate (ATP). Aside from producing ATP, the mitochrondian has other functions inside the cell such as apoptosis [Green, 1998] (programmed death of the cell), regulation of cellular metabolism [McBride et al., 2006] and heat production [Voet et al., 2016] and cell growth. Cells with high metabolic requirements can increase the number of mitchondria they contain in order to meet their needs [O'Connor et al., 2010]. For example, red blood cells have no mitochondria whereas liver cells can possess over 2000 [Voet et al., 2016] The number of mitochondria in the same type of cell can also vary. The number of mitochonria in muscle cells of those who exercise frequently are higher than those who are sedentary. The lack or misfunction of mitochondria in cells has been implicated in the development of several diseases in humans, including cardiac dysfunction [Lesnefsky et al., 2001] heart failure [Dorn et al., 2015] and autism [Frye and Rossignol, 2011].

Structurally the mitochondrion takes the form of a cylinder with rounded ends that ranges from one to many micrometers long. Like the nucleus it is separated from the cytoplasm by a double membrane. The outer membrane forms the cylinder like shape, whilst the inner membrane is delicately folded so that it

projects into the insides of the organelle [Alberts et al., 2013]. The mitchondrion retains its own genome, which is similar to the alpha-Proteobacteria prokaryotes, and reproduces by dividing into two. This is due to the fact that the mitochondrion was once a prokaryote that was engulfed by a more sophisticated eukaryote when feeding. It is thought that once eaten the prokaryotes continued to provide energy for the cell and so the host eukaryote came to rely on their presence is a symbiotic relationship. Over a very long time the internal prokaryotes lost their ability to live independently [Mai et al., 2005].

The endosymbiotic hypothesis could also explain how chloroplasts came about. These large green organelles are found only in plant and algae cells. Chloroplasts capture energy from the sun which is used to produce glucose and oxygen in a process known as photosynthesis. From a global perspective it could be argued that chloroplasts play an even more important role than the mitchondria for life on earth, due to their oxygen producing nature.

2.3.4 Other organelles

Besides the nucleus, mitochondria and chloroplasts, there are many other organelles in the cytoplasm of a eukaryote. In an averagely sized eukaryote the total volume of organelles is just less than half of the volume of the cell [Alberts et al., 2013]. The organelles in the cell perform specialised functions. Most exist to deal with the cell's need to import nutrients and export waste materials and synthesised products. In some cells specific organelles are extremely enlarged, or they may be particularly plentiful. Other important organelles in the eukaryote include the rough and smooth ER, the cytoskeleton and the peroxisome.

The ER is one of the larger organelles in the cell. It exists in two types: the rough and the smooth. The rough ER is studded with ribosomes which are responsible for the synthesis of protein. In most cells the smooth ER (lacking the ribosomes on the rough ER) is smaller than the rough ER. In the cells of the adrenal gland the smooth ER is the site of steroid hormone synthesis, in liver cells it is where alcohol is detoxified. Proteins made at the ER are packaged into

small containers called vesicles at another organelle called the Golgi apparatus. The protein containing vesicles are then sent on to the locations in the cell where they are needed.

2.4 Prokaryotes

Prokaryotes distinguish themselves from the eukaryotes by their lack of membrane bound organelles. Consisting of the bacteria and archaea domains of life they are the more diverse than the eukaryotes and they outnumber any other organisms on Earth. Some prokaryotes require oxygen in order to generate energy and some die with the slightest exposure [Alberts et al., 2013]. Prokaryotes inhabit a wide range of habitats; volcanic mud, concentrated brine, marine sediments, the sludge of sewage treatment plants and in pools beneath the frozen surface of Antarctica. Prokaryotes began to appear in their current form at a time when the Earth's atmosphere was not rich in oxygen and these harsh environments reflect this past. Some prokaryotes can draw nutrients from exotic material such as petrol or bacteria, or from completely inorganic material.

As previously stated prokaryotes do not have membrane bound organelles. Intracellular entities such as proteins, DNA and metabolites are dissolved in the cytoplasm instead of being partitioned by plasma membranes. Some bacteria do posses microcompartments [Cheng et al., 2008] which are enclosed by partially permeable protein shells [Kerfeld et al., 2010] and are thought to act as primitive organelles by enhancing or protecting key pathways within the cell [Yeates et al., 2010].

2.4.1 The Cell wall

In addition to possessing a cytoplasm and a cell membrane, which are also found in eukaryotes, prokaryotes possesses other structures. The cell wall is a common feature of the prokaryote (in addition to the plant and fungal cells) with few examples of the domain (Mycoplasma and Thermoplasma) lacking this struc-

ture. The cell wall encloses the cell membrane, increasing the cell's survivability by protecting the cell against mechanical damage and osmotic pressure. It also provides shape and rigidity [Nelson et al., 2008; Todar, 2006].

The cell wall is so important that it is used to classify one of the two domains of prokaryote, the bacteria, into two classes; Gram positive and Gram negative. The cell wall of Gram-positive bacteria is made of multiple layers of peptidoglycan whereas Gram-negative bacteria display a thinner cell wall that consists of fewer layers of the material, that is enclosed by another cell membrane [Todar, 2006]. The majority of bacteria are of the Gram negative classificiation and exhibit different responses to antibiotics when compared to Gram positive bacteria [Hugenholtz, 2002]. The sub types of bacteria are named after the famous Gram staining test that can differentiate the two strains under a microscope. As has been previously mentioned, the prokaryotic cell shape is provided by the cell wall. While bacteria come in many different shapes, most fall into three categories: spherical (cocci) rod shaped (bacilli) and spiral(spirillum and spirochet [Cabeen and Jacobs-Wagner, 2005].

These shapes are important enough to inform bacterial naming e.g. staphylococcus. In the other domain of prokaryotes, the archaea, the nature of cell is different to bacteria; this should come as no surprise given the genetic gulf between the two. In archeae there are no less than four types of cell wall. (For further information see [Todar, 2006])

2.4.2 The Nucleoid

The nucleoid is the prokaryotic analogue to the nucleus which carries the genetic material of the prokaryote. This structure is particularly large compared to the size of the prokaryote and takes on different shapes within different species. Like all structures within the prokaryote it is not bound by a membrane, rather it simply consists of molecule(s) of DNA known as a chromosomes [Madigan et al., 2010]. Along with one or more chromosome(s), which encodes genes vital for the cell's survival, most prokaryotic cells also carry one or more molecules

of DNA called plasmids. Plasmids do not form part of the cell's genome, but instead confer special behaviours upon the cell [Madigan et al., 2010; Mai et al., 2005]. The property of having a single copy of a chromosome marks prokaryotes as being genetically haploid whereas eukaryotes, typically carrying two copies of each gene, are known as genetically diploid. This distinction is important as it causes different reproductive behaviour.

2.5 Cell Geometry

Cells exist in nature in a vast range of shapes, sizes and internal organisations. In this section we review the three aspects of spatial configuration that we investigate in this thesis; differences in external cell shape, differences in internal organisation and finally processes that result in a cell changing shape.

2.5.1 Overview of cell shapes

Living cells take on a variety of different sizes and shapes leading to different surface-to-volume ratios and intracellular distances. The shape of the cell has been found to affect cellular behaviour at a structural and chemical level [Soh et al., 2010]. Cells without uniform shape have been shown to alter cellular processes in interesting ways; for example, the dendrites and soma all exhibit geometry that can lead to chemical reactions preferentially occurring in different parts of the cell [Kholodenko and Kolch, 2008; Meyers et al., 2006].

In the introduction we saw three examples of eukaryotic cells with very different shapes. These shapes help the cell in their function; the nerve cell has a long body with branching nodes at each end to reach far away locations and achieve a high level of interconnectivity, the neutrophil has no fixed shape, constantly reconfiguring itself to move and engulf pathogens, finally the red blood cell's concave hull allows for efficient gas exchange. These shapes are often compromises between the cell's differing needs. In the case of the red blood cell the *most efficient* shape for gas exchange would be a flat disc; as

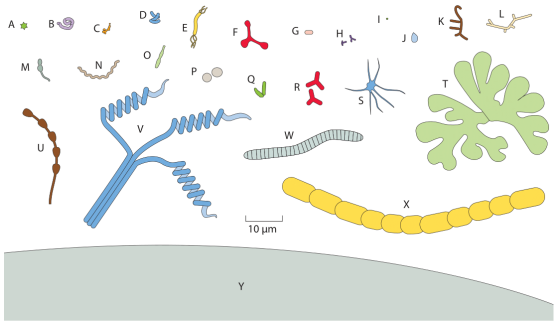


Figure 2.1: A selection of microbial sizes and shapes. These drawings are based upon microscopy images from the original literature. (A) Stella strain IFAM1312 (380); (B) *Ancylobacter flavus* (367); (C) *Bifidobacterium bifidum*; (D) *Clostridium cocleatum*; (E) *Aquaspirillum autotrophicum*; (F) *Pyroditium abyssi* (380); (G) *Escherichia coli*; (H) *Bifidobacterium sp.*; (I) transverse section of ratoon stunt-associated bacterium; (J) *Planctomyces sp.* (133); (K) *Nocardia opaca*; (L) Chain of ratoon stunt-associated bacteria; (M) *Caulobacter sp.* (380); (N) *Spirochaeta halophila*; (O) *Prosthecobacter fusiformis*; (P) *Methanogenium cariaci*; (Q) *Arthrobacter globiformis* growth cycle; (R) gram-negative *Alphaproteobacteria* from marine sponges (240); (S) *Ancalomicrobium sp.* (380); (T) *Nevskia ramosa* (133); (U) *Rhodomicrobium vanniellii*; (V) *Streptomyces sp.*; (W) *Caryophanon latum*; (X) *Calothrix sp.* (Y) A schematic of part of the giant bacterium *Thiomargarita namibiensis* (290). All images are drawn to the same scale. (Reproduced from [Milo and Philips, 2016], original credit [Young, 2006]

that the red blood cell is not a completely flat disc this would suggest that the function of gas exchange is balanced with others that depend on the shape of the cell. There is no shortage of extraordinarily shaped eukaryotes. The Photoreceptor cell is another example of a cell whose shape is adapted to its function. These retinal cells are responsible for collecting light and need to maximise their ability to do so. One adaptation of the Photoreceptor's shape is presence of the outer segment protrusion that contain photosensitive molecules. The rod subtype of photoreceptor which particularly good at detecting light has a bigger protrusion than the cone subtype.

Prokaryotes also take on a variety of shapes. Figure 2.1 shows examples of 25 different prokaryote shapes. There are a bewildering variety of shapes on display in this figure.

One of simplest measurements one can make of the cell is that of the volume

Figure 2.2: This figure illustrates the variety of shapes present in protists. Protists are drawn to scale relative to the head of a pin of approximately 1.5mm diameter. Reproduced from [Milo and Philips, 2016] (Originally from B. J. Finlay, Science 296:1061, 2002.)

of the cell. Natural cells exist in various sizes spanning seven order of magnitude from sub-micron to meter-sized. Cells from different domains vary in size. Prokaryotes are usually small, typically micron sized whereas eukaryotes can be larger, usually ranging from 5 to 50 μm with large nerve cells and egg cells at the upper end of the spectrum. It is commonly theorised that a prokaryote's small size is due to its reliance on diffusion as a method of absorbing nutrients by creating a smaller surface-area-to-volume ratio. Figure 2.2 shows the range of sizes of unicellular eukaryotic cells as compared to the head of a pin.

2.5.2 Significance of the ratio of structure in the cell

The ratio of the volume of organelles to the volume of cytoplasm in cells differs according to their function. As an example, plasma B cells have a high ratio of organelles to cytoplasm [Wintrobe and Greer, 2009]. They contain many ER and Golgi apparatus whose function is to produce and secrete antibodies. The more of these organelles that exist in the cell the more antibodies will be produced, and so in turn less volume is available for cytoplasm. The higher the ratio of organelles (cellular structures) to cytoplasm the more efficient the cell will be [Wintrobe and Greer, 2009].

Nuclear to Cytoplasm (NC) ratio is commonly used to quantify the structure

of the cell. NC ratios often vary between cell types and even within cells of the same type. An illustrative example of this diversity is the white blood cell; *Basophilic myeloblasts* typically display NC ratios of 80% whereas regular *Basophiles* have NC ratios of approximately 20%. NC ratios can also differ within the same cell at different stages in the cell's life cycle. In some cases the NC ratio can indicate the maturity of the cell; generally, as the cell develops the size of the nucleus gradually decreases in proportion to the cell. For example, the NC ratios of some "blast" forms of erythrocytes, leukocytes, and megakaryocytes steadily decreases over their life from as high as 4 to as low as 1 [Turgeon, 2005]. Abnormal NC ratios can also be predictor of malignancy [Chow et al., 2012]. The classical example of a malignant cell with a high NC ratio is the small cell carcinoma; this metric is often used in the diagnosis of many other types of tumor. NC ratio does not serve as a binary classifier for cell malignancy, rather, the further the NC ratio is from the average, the more likely the cell is to be malignant [Katsoulis et al., 1995; Montironi et al., 1990; Pienta and Coffey, 1991; Salmon et al., 1992]. Changes in NC ratio are thought to affect pathway behaviour, with some research showing the NF-kB pathway to be particularly sensitive to this spatial parameter [Ohshima et al., 2012].

In this work we extend the concept of the NC ratio to cover typical organelles and inclusions, this new metric is called the Structure to Cytoplasm (SC) ratio. Some simulations that are performed in this work use the NC ratio (when examining the size or position of the nucleus) and some simulations use the SC ratio (when examining the distribution of internal cell volume).

2.5.3 Significance of the position of the cell nucleus

The position of organelles within the cell is the second variable of internal structure that we explore. In different types of cell a given organelle can be positioned differently. This is especially true of the nucleus whose position varies (despite diagrammatic depictions to the contrary) between cell types, between cells of the same type, and within the same cell over time [Buchman and Tsai, 2008].

Nuclear position within the same cell can be contingent upon the stage in the life cycle, the migration state and differentiation status [Edens et al., 2013]. Figure 2.3 shows examples a selection of cell types in which the nucleus occurs in different positions within the same type of cell. This figure is reproduced from [Edens et al., 2013].

Nuclear positioning has been shown to be central to the regulation of some cellular processes. Conversely, it is hypothesised that irregularities in nuclear positioning may cause some diseases. Recent results from human studies and mouse models have shown that faults in mechanisms responsible for nuclear positioning are associated with disease [Gundersen and Worman, 2013]. In particular, mutations in genes responsible for nuclear lamina are associated with some genetic conditions. On the whole, however, the physiological link between nuclear position and cellular function is far from being understood [Dupin and Etienne-Manneville, 2011].

2.5.4 Significance of the distribution of internal structure in the cell

While any number of cells may exhibit the same absolute volume of organelles, the distribution of that volume may be entirely different. A particularly striking example of this disparity is displayed by the different types of fat cell.

In mammals, three forms of adipocytes exist: white, brown and beige fat cells. White fat is the most common type of fat in the body, found predominately under the skin and in the abdomen. White fat cells contain the body's main store of energy in the form of lipid droplets [Owens, 2014]. Healthy white fat cells are thought to prevent damage to organs, such as the liver, by capturing potentially toxic lipids. Brown fat cells are located in locations such as the shoulder blades and kidneys. The primary function of brown fat cells is to convert the energy stored in their glucose and lipids into heat when the organism is exposed to the cold. The third type of fat cell, the beige adipocyte, is found within deposits of white fat. It is thought to be formed from immature white fat cells when the

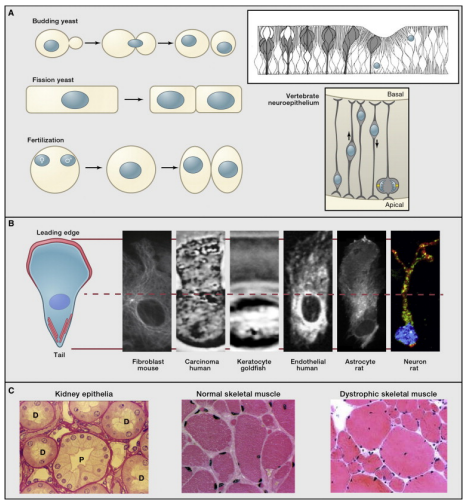


Figure 2.3: Differing nuclear positions in common cells. Schematics of nuclear positioning in dividing cells and developing epithelium. Arrows indicate movements of nuclei (blue). The nucleus is positioned relative to the plane of division in yeast and fertilized eggs. The diagram of insect optic epithelium (adapted from Patterson et al., 2004; Tomlinson and Ready, 1986) represents a longitudinal section of a larval eye disc; two nuclei are shown. Nuclei that are anterior (A) to the morphogenetic furrow (mf), which moves anteriorly, move basally. Nuclei that are posterior (P) to the furrow move apically as cells are recruited into clusters comprising ommatidium (white cells, cones; gray cells, R cells). The diagram of vertebrate neuroepithelium represents a longitudinal section of the developing cerebral cortex. Nuclei move basally during G1 and apically during G2. Mitosis (M) occurs near the apical surface. Adapted from Buchman and Tsai (2008). Figure and caption reproduced from [Edens et al., 2013]

organism is exposed to cold [Wu et al., 2012].

Structurally, brown, white and beige fat cells differ; white fat cells contain a single lipid drop whereas brown fat cells contain numerous smaller droplets and a much higher concentration of mitochondria. As well as being functionally hybrid, beige fat cells are structurally hybrid, containing a mixture of different sized lipids. Images of white, brown and beige fat cells can be seen in Figure 2.4 (figure reproduced from [Owens, 2014]). These three cells display a similar ratio of cytoplasm to organelle, but in the brown and beige cells the total volume of occupied space is less contiguous than in the white fat cell; a similar volume has been distributed differently.

Despite rich information such as this, the spatial distribution of organelles

and other intracellular inclusions is rarely utilised in the modelling of intracellular pathways [Pham and Ichikawa, 2013].

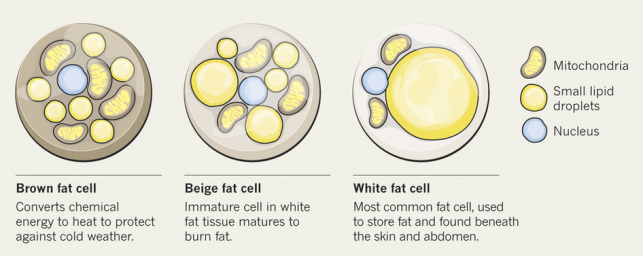


Figure 2.4: Schematic of the three different types of fat cells reproduced from [Owens, 2014]. In this diagram we see the differences in the way that a similar volume is occupied; The brown fat cell's non-cytoplasm volume is distributed between many small organelles, the white fat cell's non-cytoplasm volume is distributed over one large lipid droplet and a few small organelles and the beige fat cell is hybrid of the two.

2.5.5 Overview of processes that change the shape of the cell

Many real cells exhibit plasticity during their life cycle, changing from one shape to another in order to suit their environment [Yin et al., 2014]. Such changes can be caused by a variety of environmental factors [Young, 2006] with timescales and degrees of plasticity of varying magnitudes. For example, small shape changes can be seen in most cells due to instabilities in cytoskeletal structure [Kueh and Mitchison, 2009] with larger changes occurring in scenarios such as stem cell differentiation. Nutritional stress is another example of a shape changing factor, causing cells to undergo filamentation [Young, 2007]. In this thesis we focus on two ways in which cells can *change* shape; Hypertrophy, hypotrophy (atrophy) and filamentation.

Hypertrophy and atrophy change the size of the cell. Hypertrophy can results from both normal and abnormal processes. Hypertrophy occurs naturally in muscle cells in response to exercise. Hypertrophy of the body of many cells causes hypertrophy of the organ that it composes (in a multicellular organism). Abnormal hypertrophy can occur with the muscles of the heart, where it is

particularly associated with the left ventricle [Kusumoto, 2004]. Atrophy of the cell can be caused by either benign processes or brought about by disease or misfunction.

Eukaryotic cell shapes change when they adhere to a surface, changing their appearance to that of a "fried egg". Typically cell movement along a surface is achieved by the use of *lamellipodia*, a projection from the front edge of the cell supported by an actin mesh. When the cell moves it protrudes the front part of the cell and then retracts the rear, pulling itself along the substrate [Anderson et al., 1996]. Using such a method of movement, the cell is capable of moving at between $15\mu m$ and $20\mu m$ per minute [Holmes and Edelstein-Keshet, 2012]. Figure 2.5 shows a cartoon of the way in which these protrusions facilitate movement.

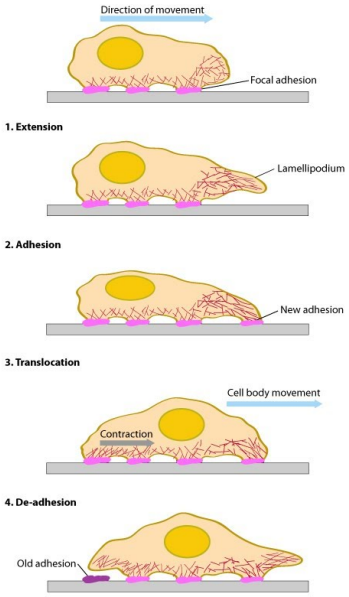


Figure 2.5: Diagram of a cell moving along a substrate. Figure reproduced from [Othmer, 2016]

Another adaptation is known as filamentation. When this process is due to stress, filamentation is caused by the cell continuing to grow whilst being unable to divide. Filamentation can also be caused by other processes, and the reasons for this are not always understood. Common theories include fluid shear

stress, anchoring themselves in soil, some forms of motility, and to predate upon Protists [Young, 2006].

2.6 Physical and chemical processes in biological cells

2.6.1 Cellular Reaction Pathways

Many chemical reactions take place in biological cells, which when taken together make up the organisms' metabolism. The behaviour of these reactions are controlled according to the cells' needs; reactions can be turned on and off, sped up and slowed down. Most cellular reactions do not take place in isolation, instead, they form part of a step by step chain called a biological pathway. In biological pathways some subset of the product from the previous stage is used as a reactant (substrate) for the next. Biological pathways are themselves organised into biochemical networks. Many chemical reactions take place within the cell - balancing the requirements of all processes is a delicate task. Cells manage reaction behaviour by moderating reaction speed. This is usually accomplished by increasing or decreasing the amount of a rate-limiting chemical. Some reaction control behaviour is particularly complex and involves reaction products that inhibit their own production in a feedback loop [O'Connor et al., 2010]. Enzymes are critical to reactions within the cell; many biochemical reactions would be extremely slow at the temperatures within the cell without their presence [Alberts et al., 2013]. Enzymes act as catalysts, increasing the likelihood of a reaction without the need for higher temperatures without being consumed. Reaction rates can be altered by the presence of inhibitor or activator molecules. Under specific conditions these are capable of binding to enzymes to prevent or enable a reaction. The situation is further complicated by the fact that control for a single pathway is shared between a number of enzymes.

The second law of thermodynamics states that a chemical reaction can only

occur if it results in a net increase in a closed system's entropy. Entropy can be expressed in terms of the change of free energy, δG, when a change in the system occurs - in energetically favourable reactions δG is positive and in unfavourable reactions δG is negative. One might look at the reactions that synthesise and maintain the highly ordered environment of the cell and immediately exclaim that thermodynamics has been violated. This concern can be assuaged by realising that biological cells are not closed systems; they take energy from external nutrients or photons from the sun, which is then used to create order within the cell [Alberts et al., 1995]. Reactions within the cell liberate energy (in the form of heat) which creates disorder outside of the cell, balancing the increase in order within.

As discussed, chemical reactions within cells are organised into biological pathways. Some pathways process signals related to changes in the environment of which the cell is located, such as infection, stress, injury, or nutrient levels. The effects of these interactions vary widely, but can include the production of a new chemical species, such as a fat or a protein, or generating cell movement. Most cellular pathways fall into one of two categories: metabolic, or signal transduction. Metabolic pathways involve the transformation of one species into another, either to be used immediately, to be stored, or to initiate other metabolic pathways. Metabolic pathways are divided into catabolic pathways (which liberate energy from nutrients) and anabolic pathways (synthesis of more complex molecules using the energy liberated from catabolic reactions). Signal transduction pathways amplify extracellular signals received at surface receptors and propagate them through the cell. Cellular behaviour is also controlled by gene regulatory networks, which are responsible for turning genes on and off. This switching governs the production of proteins. It has become clear that these reaction pathways are spatially organized within the cell, with many reactions occurring only in specific regions [Srere, 2000]. For example, within signal transduction pathways, spatial gradients and microdomains of signalling occur due to localised chemical species, such as phosphatases.

2.6.2 Common Types of Reaction

Chemists have grouped reactions into a few general groups. In this subsection, we review common reaction types that take place in the cell.

Synthesis or addition reactions are those in which two molecules are combined to form another molecule. In general we consider reactions such as:

$$A + B \longrightarrow C$$

In reality, the number of substrates and products changes. One example of this family of reactions are the set of hydration reactions. Carbon dioxide, a by-product of cellular respiration, is released into the blood stream for transport to the lungs so that it can be exhaled. In the blood stream, carbon dioxide interacts with water via an addition reaction forming carbonic acid.

$$CO_2 + H_2O \longrightarrow CO(OH)_2$$

Decomposition reactions can be seen as the opposite of synthesis reactions. It is a reaction in which a chemical compound is broken into simpler compounds or elements. The stability of a compound can be weakened by the presence of conditions such as heat or a solvent. In general, the reaction can be represented as:

$$AB \longrightarrow A + B$$

Decomposition reactions are present in biochemistry. One example of this reaction is the disposal of hydrogen peroxide. Cells produce the enzyme catalase to remove this harmful by-product of respiration. Chemically the decomposition of hydrogen peroxide can be described as:

$$2H_2O_2 \longrightarrow 2H_2O + O_2$$

The word redox is a portmanteau of reduction and oxidation. These two

words describe properties of electron transfer in chemical reactions. Oxidation is defined as the loss of electrons from a molecule, atom or ion. Reduction is defined as the gain of electrons by a molecule, atom or ion. Redox reactions are those in which both of these processes occur at once and electrons are transferred between chemical species.

Many vital cellular processes involve redox reactions. Cellular respiration, the process by which mitochondria liberate energy from nutrients, is such a reaction.

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O$$

In this reaction, the glucose is oxidised to CO_2 and oxygen is reduced to water. Photosynthesis uses sunlight energy to produce glucose. Carbon dioxide undergoes reduction to glucose and water is oxidised to oxygen.

$$6CO_2 + 6H_2O \longrightarrow C_6H_{12}O_6 + 6O_2$$

Many systems within the cell are understood to be controlled by oscillating reaction systems [Vîlcu and Bala, 2004] with circadian rhythms [Noyes and Field, 1974], insulin and hormone secretions being well known examples. There are also oscillating mechanisms that are much faster than this [Novák and Tyson, 2008]. Specific examples include mitosis control in slime mold [Tyson and Kauffman, 1975] and glycolysis within yeast cells [Hess et al., 1973]. The presence of a cell membrane may induce oscillatory behaviour in systems that would otherwise not display this behaviour; [Hahn et al., 1973] showed that coupling membrane permeability with reaction causes such behaviour. Another possible explanation for oscillatory behaviour is given in [Hess and Boiteux, 1971]; the authors propose that activation and inhibition of enzymes may be responsible. As activation/inhibition of enzymes for the control of reaction speed is common within cells [O'Connor et al., 2010] oscillatory behaviour may be present in many other cell pathways in many circumstances, especially within feedback

loops.

2.6.3 Transport Mechanisms in Cells

The theory of diffusion is a fundamental process of the natural world. It has been used to model processes in many disciplines, including but no limited to: economics, physics, chemistry, and most importantly for this thesis, biology. "Diffusion is the process by which matter is transported from one part of a system to another as a result of random molecular motions down a concentration gradient" [Crank, 1979]. In the classical experiment that demonstrates diffusion, a small volume of iodine is placed at the bottom of a tall cylindrical container. Water is then slowly and carefully poured on top of the solution as to minimise the mixing effects of pouring. At the start of the experiment the boundary between the iodine and the water is clearly demarcated, but after some time the iodine spreads upwards through the water, with the top of the container having the lightest shade, and the bottom the darkest shade of purple. After sufficient time has passed, the entire cylinder contains uniformly coloured liquid that is a lighter shade than the pure iodine present at the start of the experiment. The iodine has diffused through the water and the iodine/water solution is now well mixed.

If it were possible to track the motion of an individual molecule of iodine under a theoretical microscope then one would see an entity in constant translational motion, bouncing off the molecules of water, moving randomly with no preference for any particular direction of motion. This track of movement, if plotted, would appear as the familiar "random walk" pattern of movement. In this regime, it is possible to calculate the mean-square displacement that a molecule will travel in a particular time but it is impossible to predict the trajectory of the molecule. This type of molecular movement is know as Brownian Motion (BM), after Robert Brown who observed the random walk of a pollen particle in water under a microscope.

If BM is random, how do molecules end up moving to less concentrated

regions? Consider a thin horizontal slice of the container, dv_c, containing a mixture of iodine and water. In addition, consider identical slices, also containing an iodine/water mixture, above - dv_a, and below - dv_b the central slice - dv_c. Whilst it is impossible to say which water/iodine molecules will randomly move from dv_a to dv_c or from dv_b to dv_c, it is possible to calculate what fraction will in a given time period. As there are more iodine molecules in dv_b than in dv_a the net transfer of iodine molecules will be into the central slice, dv_c. Likewise in slice dv_c there is an equal chance of a molecule moving into dv_a or dv_b, but as there are more iodine molecules in dv_b the net movement of molecules will be into dv_a. Diffusion is the macroscopic consequence of Brownian motion.

It was Adolf Fick who in 1855 quantified diffusion [Fick, 1855] by casting particle diffusion in the same vein as heat conduction [Fourier, 1822] derived by Fourier in 1822. Fick's law states that the diffusion of a substance, within isotropic media, through unit cross sectional area is proportional to the concentration gradient:

$$F = -D\frac{\partial C}{\partial x} \tag{2.1}$$

F is the rate of transfer per unit area, C is the Concentration of the diffusing substance, x is the length perpendicular to the cross sectional area and D is the diffusion coefficient. The value of D, measured in cm^2s^-1 , is usually constant in solution, but may vary in polymers as a function of the concentration. The negative sign denotes that diffusion occurs in the opposite direction to increasing concentration [Crank, 1979].

This is the intuition of how Brownian motion results in the diffusion process; now explore the quantitative explanation. This derivation appears in [Bressloff, 2014]. Let us first consider a random walker on a one dimensional lattice.

Let a walker begin at $x = 0$ and take steps of length Δx. At each timestep, Δt, the walker has an equal probability of moving left or right. The direction of each step is independent of the previous. The displacement at each step, s_i,

is given by:

$$s_i = \begin{cases} +\Delta x, \ with\ 0.5\ probability \\ -\Delta x, \ with\ 0.5\ probability \end{cases} \tag{2.2}$$

After N timesteps the position (and displacement, x) of the walker is

$$x(N) = \sum_{i=1}^{N} s_i \tag{2.3}$$

As the walker moves both left and right with equal probability, it is obvious that:

$$\mathbf{E}(x(N)) = 0 \tag{2.4}$$

Equation (2.4) tells us that the average position of the walker will be at its origin. It is very important to realise that this does not mean that the particle will *always* be positioned at the origin. Rather, it means that the probability distribution of displacement is *centered* at the origin. As the number of steps, N, increases the probability distribution widens.

As the lattice spacing, Δx and the timestep size, Δt tend to zero, the system can be rewritten as a diffusion equation. We define a function $P(x,t)$ that gives the probability that a walker is at a particular lattice position, x, after N steps:

$$P(x_0, t_0 + \Delta t) = \frac{1}{2}P(x_0 + \Delta x, t_0) + \frac{1}{2}P(x_0 - \Delta x, t_0) \tag{2.5}$$

We can then use the Taylor expansion on the function, $P(x,t)$:

$$P(x_0, t_0 + \Delta t) = P(x_0, t_0) + \frac{\partial P}{\partial t}(x_0, t_0)\Delta t + \frac{1}{2}\frac{\partial^2 P}{\partial t^2}(x_0, t_0)(\Delta t)^2 + ... \tag{2.6}$$

$$\frac{1}{2}P(x_0 + \Delta x, t_0) = \frac{1}{2}P(x_0, t_0) + \frac{1}{2}\frac{\partial P}{\partial t}(x_0, t_0)(\Delta x) + \frac{1}{4}\frac{\partial^2 P}{\partial t^2}(x_0, t_0)(\Delta x)^2 + ... \tag{2.7}$$

$$\frac{1}{2}P(x_0-\Delta x,t_0)=\frac{1}{2}P(x_0,t_0)+\frac{1}{2}\frac{\partial P}{\partial t}(x_0,t_0)(-\Delta x)+\frac{1}{4}\frac{\partial^2 P}{\partial t^2}(x_0,t_0)(-\Delta x)^2+... \tag{2.8}$$

We then substitute (2.7) and (2.8) into (2.6) and obtain:

$$\frac{\partial P}{\partial t}(x_0,t_0)\Delta t+...=\frac{1}{2}\frac{\partial^2 P}{\partial t^2}(x_0,t_0)(\Delta x)^2+... \tag{2.9}$$

As both Δx and Δt are small, the higher order terms in the Taylor expansion can be dropped. Dividing (2.9) by Δt:

$$\frac{\partial P}{\partial t}(x_0,t_0)\approx\frac{(\Delta x)^2}{2\Delta t}\frac{\partial^2 P}{\partial t^2}(x_0,t_0) \tag{2.10}$$

In the limit $\Delta x\to 0$, $\Delta t\to 0$ and where $\frac{(\Delta x)^2}{\Delta t}$ is still finite this becomes an exact relation. Further if we define:

$$D\equiv\frac{(\Delta x)^2}{2\Delta t} \tag{2.11}$$

We get the famous diffusion equation:

$$\frac{\partial P}{\partial t}(x_0,t_0)=D\frac{\partial^2 P}{\partial t^2}(x_0,t_0) \tag{2.12}$$

Remarkably, starting with a stochastic system had yielded a deterministic approximation.

As with iodine in water, biologically relevant molecules are in constant motion within the cell, colliding with other molecules of the highly packed cytoplasm. Diffusion provides an efficient way for a molecule to explore the inside of a cell [Alberts et al., 1995]. Many cells take advantage of diffusion's low energy requirement for molecular transport [Soh et al., 2010] (where chemical concentrations are particularly high) by having not evolved complex transport mechanisms.

In 2.6.1, we saw how pathways are highly influential in the behaviour of cells,

with signalling pathways responsible for information exchange and processing and metabolic pathways constituting the cell's metabolism. The reactions that, in concert, form these pathways could not take place if the necessary reactants were not in the correct location. It has been demonstrated in a number of cells that this required transport of molecules is accomplished through diffusion [Soh et al., 2010]. Both metabolic and signalling pathways use diffusion as a transport mechanism [Li et al., 2015]. Glycolysis and lipolysis (metabolic pathways), and MAPK (a signalling pathway) all rely on diffusion as a means of transport across the cytoplasm. Any cellular processes that involve the shuttling of molecules between the cytoplasm and the nucleus also rely on the same process. It is important to note that when the diffusion process is much faster than the reaction process, the effect of diffusion on the dynamics of the system is decreased. In these cases, reaction-only models may be sufficient. In this thesis we therefore limit discussions of systems to those in which *both* reaction and diffusion have a significant impact on the dynamics of the system.

Aside from serving as a transport mechanism for many biological pathways diffusion is, particularly in bacterial cells, responsible for the delivery of nutrients and macromolecules through the cell membrane [Young, 2006]. [Beveridge, 1988] commented that "diffusion is a prime factor for bacterial life and that the wall, by determining shape, will dictate diffusion efficiency". Diffusion is commonly cited as the reason that prokaryotes are so small - the rate that nutrients can be absorbed at is a function of surface area [Young, 2007]. In addition, as the surface area to volume ratio increases, the cell becomes smaller and the amount of cytoplasm that each unit area of membrane has to support becomes smaller. However, diffusion alone cannot be responsible for cell shape and size. As bacterial shape and size occupy such an enormous range, from the diminutive *Pelagibacter ubique*, occupying 0.01 μm^3 [Rappé et al., 2002] to *Epulopiscium fishelsoni* occupying $10^8 \mu m^3$[Schulz and Jørgensen, 2001], diffusion cannot be the only factor that influences cell size.

2.6.4 Determining diffusion coefficients

The quantity, D, in (2.12), is known as the diffusion coefficient - it is the constant of proportionality between the rate at which the chemical is transported in the medium to the concentration gradient of the chemical. Inferring the value of these diffusion coefficients is of central importance for simulation in cell biology - models of reaction and diffusion are not "magic" and still require parameters. Any method that measures concentration profiles over time can be used to determine additional diffusion coefficients. Here we describe three common techniques for measuring this value: Fluorescence Recovery after Photobleaching (FRAP), Fluorescence Correlation Spectroscopy (FCS) and Single Particle Tracking (SPT).

The FRAP microscopy technique was first invented in the 1970s [Axelrod et al., 1976] [Peters et al., 1974]. In the FRAP method of diffusion coefficient determination a high-intensity laser is shone at the cell causing photobleaching of a micrometer area. As non-bleached molecules diffuse into the bleached area the total fluorescence recovers in the bleached spot. The increasing fluorescence of this area over time is measured, plotted against time and finally a mathematical FRAP model is fitted to the curve. The diffusion constant is extracted from the curve.

FCS, a complementary technique to FRAP, was developed in approximately the same era [Ehrenberg and Rigler, 1974] [Magde et al., 1974]. The intensity of fluorescence is measured at one sub micrometer sized area in the cell with extremely sensitive avalanche photo diode detectors. When a fluorescent molecules enter or leave the small measured area the total fluorescence of that area increases or decreases respectively. The results from this experiment take the form of fluorescence time series. When autocorrelation analysis is applied to the time series the chemical's diffusion coefficient and concentration can be extracted [Gösch and Rigler, 2005] [Haustein and Schwille, 2004] [Krichevsky and Bonnet, 2002]. FCS is suited to extremely fast processes.

Measurements from FRAP and FCS experiments are gained from the move-

ment of many individual molecules. In contrast, SPT is capable of tracking single molecules, or nanoparticles directly (Saxton and Jacobson 1997; Suh et al. 2005). This is accomplished using sophisticated electron multiplying Charge-Coupled Device (CCD) cameras in conjuction with widefield laser excitation. Molecules and nanoparticles are visible as points of light with a Gaussian intensity distribution. Despite limitations in optical resolution, the positions of molecules can be calculated with nanometer precision [Kubitscheck et al., 2000] [Thompson et al., 2002]. Tracking of the points of light is provided by image recognition algorithms [Sbalzarini and Koumoutsakos, 2005] [Anthony et al., 2006]. Once the positions of the molecules have been determined trajectories can be determined and diffusion coefficients calculated.

2.7 Dynamical systems within cells

In their total, reaction and diffusion account for a large proportion of the cell's behaviour or *dynamics*. When reactive and diffusive systems act in concert they are known as *Reaction-Diffusion systems*. All work in this thesis is concerned with modelling cellular behaviour as reaction-diffusion systems. Reaction-diffusion systems (sometimes known as a reaction-diffusion network) are a subset of dynamical systems. Before further exploring the nature of reaction-diffusion systems we look at what defines the family of systems from which reaction-diffusion descends.

Cellular dynamics is the study of dynamical systems in cells. A dynamical system is a real world system that can be modelled as possessing a set of variables that evolve over time according to a set of rules [Arrowsmith and Place, 1990; Brin and Stuck, 2002; Broer and Takens, 2010; Irwin, 2001; Jost, 2006; Katok and Hasselblatt, 1997; Ott, 2002]. Familiar examples include the position of a pendulum, the velocities of the planets in the solar system or the number of gorillas in a jungle. Two predominant classes of dynamical system exist, those that progress in discrete time and those that progress in continuous time. More

formally, we can write down three conditions that must be satisfied for a system to be considered dynamical [Katok and Hasselblatt, 1997]:

1. A state space, the elements of which describe the possible states of the system.

2. A time regime; either discrete or continuous. The time regime may be permitted to only extend into the future (an irreversible dynamical system) or into the past (known as a reversible dynamical system).

3. A time evolution law

When applied to cell simulation, the state space represents the concentration of all chemicals at all points in the virtual cell and the time evolution law represents our model of reactive and diffusive processes in the cell.

The time evolution laws of dynamical systems are naturally modelled by differential equations, which we derive from difference equations. If f represents the time evolution law, then the evolution of the quantity of interest, x, is given by:

$$x_{n+1} = f(x_n) \tag{2.13}$$

Equation 2.13 shows that the value of x at the next point in time is equal to a function of its value at the previous point in time. Equation 2.13 can be rewritten as a difference equation:

$$x_{n+1} - x_n = f(x_n) - x_n \tag{2.14}$$

Then if we define:

$$g(x) \equiv f(x) - x \tag{2.15}$$

we can rewrite Equation 2.14 as:

$$x_{n+1} - x_n = g(x_n) \tag{2.16}$$

Equation 2.16 can be read as "if we change the system by one timestep then x will change by $g(x)$. Equation 2.16 is the discrete analogue of a differential equation, and in the continuous regime we can rewrite Equation 2.16 as:

$$\frac{dx}{dt} = g(x) \tag{2.17}$$

Equation 2.17 is reproduced from Wolfram mathworld [2]. Models of cellular dynamics tell us how behaviour (commonly concentration) at locations within the cell vary over time according to an iteratively applied time evolution law (reaction and diffusion) that is informed by observed or theorised cell behaviour. In the next section we explore the time evolution laws that are commonly used in simulations of cellular dynamics, namely reaction-diffusion models.

2.8 Cellular dynamics are spatially sensitive

It is becoming increasingly clear that biological pathways are spatially organised within the cell, with some experiments suggesting that portions of the reaction sequence occur preferentially within specific volumes in the cell [Kholodenko, 2009; Srere, 2000]. Some models of biological pathways make the implicit assumption that chemical concentrations are uniform across the internal volume of the cell [den Breems et al., 2014; Gong et al., 2010; Orton et al., 2005]. Chemical concentrations are not uniform within the cell. This can become the case when reaction volumes become relatively large and chemically connected to other cellular processes - pathway chemical concentrations may become heterogeneous. The heterogeneity of pathway chemical concentrations can form concentration gradients sometimes called *microdomains*. It should be noted, that whilst non-spatial models cannot display microdomains/concentration gradients they are

[2] reproduced from http://mathworld.wolfram.com/DynamicalSystem.html

able to answer other research questions that do not depend on spatial structure.

It has been experimentally observed that microdomains occur in a variety of cells and pathways [Howard, 2006]. Common examples include phosphorylated stathmin/oncoprotein 18 [Niethammer et al., 2004], MAPK Fus3 [Maeder et al., 2007], and Aurora B kinase [Fuller et al., 2008]. Most interestingly there is increasing evidence that without these concentration gradients some biological processes and pathways would not be able to take their experimentally recorded form [Meyers, 2012]. For example, the propagation of an action potential along the axon could not take place without a regulated set of concentration gradients. The propagation of calcium waves, that are part of the continuation of development of an egg after fertilization, are sensitive to calcium concentration gradients and the positioning of release channels [Chen et al., 2014].

The cause of these concentration gradients is an important field of study in cell biology, but it is becoming clear that the shape, size and position of substructures within the cell all play a pivotal role in inducing them and thus affect pathway behaviour [Meyers et al., 2006; Neves et al., 2008; Pham and Ichikawa, 2013]. Examples of this behaviour have been documented in a number of experimental and computational studies which we cover below.

2.8.1 Cell size

Signalling pathways are also thought to be sensitive to cell size; signalling components are more likely to be deactivated the further that they travel within the cell towards their target [Meyers et al., 2006]. Further evidence of the sensitivity of signalling behaviour to geometry can be seen within computational models of yeast cells, with variable geometry producing differing dose/response behaviour within MAPK cascades [Zhao et al., 2011]. The cyclic adenosine monophosphate (cAMP) messenger participate in many signalling pathways including the regulation of vital processes such as glycolysis and lipolysis [Alberts et al., 2013]. Simulations of this pathways reveal that in model cells of smaller diameters, cAMP concentrations are higher than those in larger cells, such as

neural soma. This is due to the progressive hydrolisation of messenger molecule as the distance into the cell's cytoplasm increases.

2.8.2 Cell shape

Cell shape is also responsible for controlling the formation of intracellular concentration gradients [Cowan et al., 2012; Neves and Iyengar, 2009; Neves et al., 2008]. Signalling chemicals such as CDC42 have been shown to become preferentially activated in spatial environments with larger surface-to-volume ratios. cAMP concentrations also achieve higher concentrations when the migrating leading edge of the cell is thinner than the cell body [Kritikou, 2008]. The NF-kB signalling pathway is another thoroughly studied example of a spatially sensitive pathway [Dolcet et al., 2005; Terry and Chaplain, 2011]. NF-kB is a transcription factor that is involved in cell's regulatory response to stress, immune and inflammatory respsonses. Its misregulation is strongly associated with many different kinds of cancer [Weinberg, 2013]. Numerical simulations demonstrate the dampening effect that elongating the cell has on the oscilliatory behaviour of NF-kB [Terry and Chaplain, 2011]. Further, we show in [Sayyid and Kalvala, 2016] that common bacterial cell shapes alter the concentration gradients in model oscillatory cellular pathways.

2.8.3 Organelle Position

Organelle positioning has also been implicated in the generation of concentration gradients. This can occur by two mechanisms; chemical components can be spatially restricted from the rest of cytoplasm and then released or the presence of the organelle could physically disrupt the pattern formed by the reaction-diffusion network. The location of endosomes and the Golgi apparatus play a critical role in regulating signal transmission to the nucleus [Hwang et al., 2014], with endosome positioning implicated in neuron polarisation, growth and signal moderation [Eva et al., 2010, 2012; Golachowska et al., 2010; Higuchi et al., 2014; Sadowski et al., 2009]. Mitochondrial location has also been shown to affect the

pathway behaviour responsible for the formation of the axon and operation of the synapse [Courchet et al., 2013; Sheng and Cai, 2012; Spillane et al., 2013]. The effect of ribosomal location on pathway progression has also been modelled. Spatially sensitive models exist for the Hes1 [Sturrock et al., 2011a], p53 [Gordon et al., 2009; Sturrock et al., 2011b], and Notch-Delta signalling [Terry et al., 2011] pathways. The response of the NF-kB pathway to the presence of the larger organelle has also been computationally explored. The NF-kB pathway undergoes abnormal activation in malignant cells, stimulating the growth of more malignant cells [Inoue et al., 2007; Kim et al., 2006; Moss et al., 2008]. As previously stated NC ratios also take on abnormal values in malignant cells. Thus the interaction between NF-kB and NC ratio has been computationally studied; different sized nuclei alter the oscillatory behaviour of the pathway [Ohshima et al., 2012; Terry and Chaplain, 2011].

2.8.4 Kinetic Parameters

Gradients of concentrations may also be caused by reaction locality, such as the activation of a chemical at the cell membrane upon entering the cell and subsequent deactivation at the nucleus. In this situation even without the activation/deactivation mechanism a gradient would be formed as the chemical diffuses from a higher concentration at the membrane to a lower concentration at the nucleus. In this case, over time as the concentrations equalised, this gradient would fall to zero and no additional transport would occur. Gradients are also displayed when chemicals of differing diffusivity constants participate in the same reaction; for example, proteins in their active and inactive forms often diffuse at different rates [Kholodenko, 2009]. Concentration gradients, however they are formed, change reaction system dynamics; non-spatial models are inherently unable to capture this behaviour and so may predict less accurately.

The effect of redistribution of internal volume, taking the form of organelles or inclusions, on cellular pathways remains poorly studied [Pham and Ichikawa, 2013]. We hope to refine our knowledge of this field with the simulations per-

formed in this work.

2.8.5 Simulation can explore pathway responses to cell organisation

Whilst the work described in this section has been successful in elucidating the link between configuration and function, in general defining this link is non-trivial and remains a challenge in cell biology. This is due in part to the difficulty of constructing experiments in which substructure configuration can be controlled. Recent research [van Bergeijk et al., 2015] has seen the development of techniques for substructure control, specifically organelle position *in vivo*. It is hoped that spatio-temporal simulations will become the norm, rather than the exception for the simulation of pathway dynamics. Many examples of research within the field suggests that spatial simulations are *vital* for improving accuracy [Mazel et al., 2009; Parulek et al., 2009].

2.9 Reaction-Diffusion models of cell dynamics

2.9.1 What are reaction-diffusion models?

Reaction-diffusion models are mathematical models that capture the reactive and diffusive behaviour of physical systems and so are able to predict species concentrations in time and space. These models are most naturally applied to chemical systems, with the species in the model being chemical species.

Reaction-diffusion models use in a variety of disciplines. They are applied in ecology, to predict population sizes [Cantrell and Cosner, 2004; Cosner, 2008], in geology to model sandbank geometry and mineral deposit distribution [Ball, 2015; Bischoff and Piper, 2013], in physics to model neutron diffusion and the formation of spiral galaxies [Leung, 1988; Wilhelmsson and Lazzaro, 2000], in epidemiology to model the spread of diseases such as rabies [Allen et al., 2008; Brauer et al., 2001] and in economics and management [Bokhari and Sadek,

2011].

Aside from predicting chemical quantities, reaction-diffusion has been found to model biological pattern formation. This phenomenon was first explored in *The Chemical Basis for Morphogenesis* [Turing, 1990] in which Alan Turing showed that a system of reacting and diffusing chemicals can evolve from a uniform state to a complex patterned result, taking the form of naturally occurring stripes, spots and spirals. This is counter intuitive to the outcome that one might expect, that is either a new uniform result, or chaos. Turing's model (also known as the Turing system) is one of the most general and versatile mechanism by which these patterns are created. Others of similar power include the Gray-Scott system which is also capable of mimicking biological pattern formation [Maini et al., 1997]. Figure 2.6 shows how well an artificially generated Turing pattern (on the right) mimics the camouflage of a fish.

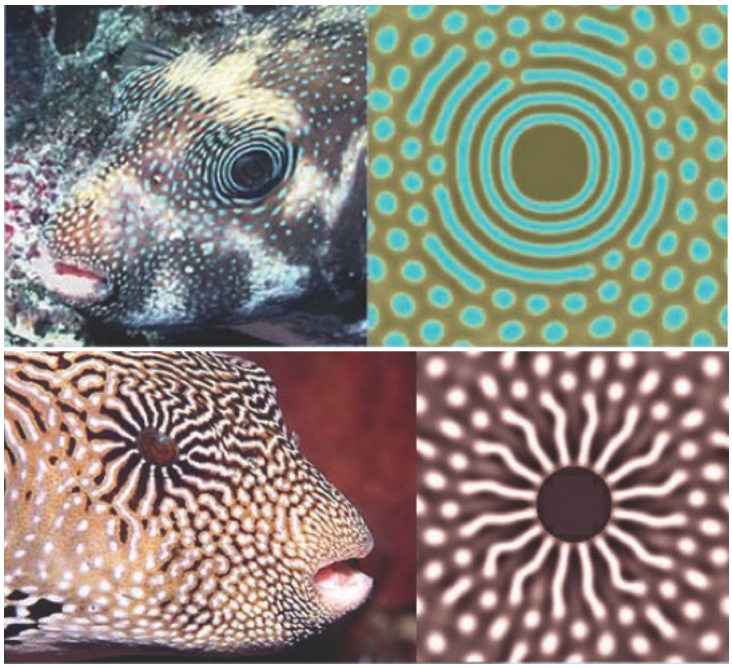


Figure 2.6: Artificially generated Turing pattern (on the right) mimics the camouflage of a fish. Credit: Fish by Massimo Boyer; simulations from [Sanderson et al., 2006]

One property of the Turing system crucial to its success in describing natural

phenomena is that of a reactant being able to make more of itself [Ball, 2015]. This property is formally known as auto-catalysis, acting as a catalyst for its own creation in a positive feedback loop. "The more that is made, the faster it appears". In this kind of reaction diffusing system new species are created at every point the other species moves into filling the system. This is a well known property of many oscillatory chemical reactions including the Lotka-Volterra (LV) system (for an explanation see section 4.2.2) that we use as a representation for this class of reactions in this thesis.

Not all reaction-diffusion systems produce well organised patterns. Some evolve into a steady state, but in some situations, reaction-diffusion systems can produce organised travelling groups of concentration known as wavefronts of concentration. Their existence within biological systems is well studied [Wang et al., 2006]. Such waves of concentration are found in the simulation results of this thesis, further confirming the choice of reaction systems as biologically appropriate.

Reaction-diffusion models have many possible formulations and these are discussed at great length in Chapter 3. For now, we present the classical differential equation model of reaction-diffusion:

$$\frac{\partial u}{\partial t} = D\nabla^2 u + R(u) \tag{2.18}$$

or alternatively:

$$\frac{\partial u}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + R(u) \tag{2.19}$$

Where, u is the concentration, D is the diffusion coefficient and R accounts for a system of reactions described by a set of Ordinary Differential Equation (ODE)s. Solving this system gives a discretised grid of concentration values that vary over time. To do this we need to solve both Partial Differential Equation (PDE)s (∇u) and ODEs ($R(u)$). In real cellular systems there can be many different chemical species, meaning that the above equation has to be

solved for each species. This can become very computationally expensive.

In all reaction-diffusion simulations in this thesis we use Neumann boundary conditions. Using a Neumann boundary condition means that we can set the value that the simulation takes when encountering an obstacle (such as the edge of the cell or an organelle) in the simulation. We set boundary cells, that would be included in the computational stencil if they were not in a boundary, as having the same concentration as the reacting/diffusing cell. This means that no concentration can leave the system. Boundary conditions were verified by running the simulation for many timesteps using a diffusion only systems and the Bi-Molecular Reaction (BMR) system. Conservation of concentration was verified.

2.10 Summary

This chapter has provided the biological background required to read this thesis. Cell anatomy, cellular dynamics and representative models have been discussed. The modelling of cellular dynamics is given a thorough treatment in chapters 3 and 4. Related work in cell dynamics simulation along with its key findings have been presented. Most importantly, Section 2.5 provides the reader with the biological motivation for performing the simulations in chapters 5, 6 and 7 which are central to this thesis.

CHAPTER 3

Reaction-Diffusion Models of Cellular Dynamics

3.1 Mathematical Models of Reaction

In this chapter we derive numerical methods for the simulation of reactive and diffusive processes. Reaction models that ignore spatial information have been well studied. These models assume the system is *well-stirred*, meaning that the species are evenly distributed throughout the reaction vessel. Derivations in this chapter are established, some derivations are taken from specific works and these have been explicitly cited. Where no individual work has been cited, the derivations are adapted and combined from [Higham, 2001, 2008; Ilie et al., 2009], which do not themselves always tie a particular derivation to an individual piece of work.

3.1.1 Stochastic chemical kinetics

Consider a well-stirred system containing N chemical species, $S_1, S_2, ..., S_N$ that may take part in one or more of M types of chemical reactions $R_1, R_2, ..., R_M$. We assume that the system is in thermal equilibrium, with a constant volume, V. The number of molecules of each species are known at time $t = 0$. The state of the system at time t is stored in the *state vector*,

$$\boldsymbol{X}(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ \vdots \\ X_N(t) \end{pmatrix} \tag{3.1}$$

where $X_i(t)$ is the number of S_i molecules at time t [Higham, 2008]. The state vector is updated every time a reaction occurs. This update process is a discrete Markov process. We seek to find the state vector, $\boldsymbol{X}(t)$, given we know the initial state of the system: $\boldsymbol{X}(t_0) = x_0$. Each reaction R_j alters the state of the system according to that reaction's state change vector, $\boldsymbol{\nu}$:

$$\boldsymbol{\nu}_j = \begin{pmatrix} \nu_{1j} \\ \nu_{2j} \\ \nu_{3j} \\ \vdots \\ \nu_{Nj} \end{pmatrix} \tag{3.2}$$

ν_{ij} is the change in the number of S_i caused by one application of reaction R_j. When the state change vectors for all reactions are considered they can be written in terms of the *stoichiometric matrix*, σ. $\sigma = \nu_{ij}$ where $1 \leq i \leq N$ and $1 \leq j \leq M$. A single reaction, R_j can be thought of in terms of how likely it is to occur. For each reaction we can write down a *propensity function*, a_j which tells us the probability of a single reaction event R_j occurring on the time interval $[t, t + dt]$. In the area of chemical kinetics this is distinct from a probability as it tells us the *time interval* over which the reaction can occur.

$$p(R_j) = a_j(\boldsymbol{x})dt \tag{3.3}$$

.

Let us now consider some common reactions and try to derive their propensity functions. For a first order reaction such as:

$$\mathrm{S_a} \xrightarrow{\mathrm{c_j}} \text{products}$$

the propensity function takes the form $a_j(\boldsymbol{x}) = c_j x_a$. That is, the probability of this reaction occurring is proportional to the number of molecules of species S_a (denoted by x_a) and the constant of proportionality is the rate constant, c_j. For a second order reaction such as

$$\mathrm{S_a} + \mathrm{S_b} \xrightarrow{\mathrm{c_j}} \text{products}$$

it can be easily seen that the propensity function takes the form $a_j(\boldsymbol{x}) = c_j x_a x_b$.

These propensity functions are derived from approximation in statistical physics and as such a derivation is outside of the scope of this thesis. Interested readers are directed to [Gillespie, 1992]. Further approximations of this sort result in the reaction rate equation, which is discussed later.

3.1.2 The chemical master equation

The most accurate representation of the kinetics of a well-stirred reaction system is that of the Chemical Master Equation (CME) which we explore in this subsection. The probabilistic nature of Equation 3.3 means that it is impossible to predict the exact nature of $\boldsymbol{X}(t)$. However we can think about the probability that the state vector at time t is in any given state ($\boldsymbol{X}(t) = \boldsymbol{x}$) if we know $\boldsymbol{X_0}(t) = \boldsymbol{x_0}$. This idea is expressed in Equation 3.4.

$$P(\boldsymbol{x}, t \mid \boldsymbol{x_0}, t_0) = Prob\{\boldsymbol{X}(t) = \boldsymbol{x}, \, given \, \boldsymbol{X}(t_0) = \boldsymbol{x_0})\} \tag{3.4}$$

To compute the probability in Equation 3.4 we begin by considering the probability $P(\boldsymbol{x}, t + dt \mid \boldsymbol{x_0}, t_0)$. dt is considered to be small enough such that only one reaction can occur on the interval $[t, t + dt)$. At time $t + dt$ the system will be in state $\boldsymbol{x}$ if either of the following events occur:

1. The system already occupied this state at time t and hence no reaction occurred

2. A reaction, R_j $1 \leq j \leq M$ from the list of reactions occurred. If this was the case then the system was in a state $\boldsymbol{x} - \boldsymbol{\nu_j}$ at time t and the reaction R_j occurred in the time interval $[t, t + dt)$.

Summing these outcomes together, it is apparent that there are $M + 1$ mutually exclusive events that lead the system to being in state $\boldsymbol{x}$ at time $t + dt$.

The probability that reaction R_j occurred in the interval $[t, t+dt)$ when the system was in the state $\boldsymbol{x} - \nu_j$ is $a_j(\boldsymbol{x} - \boldsymbol{\nu_j})dt$. The probability that no reaction occurred in the interval $[t, t+dt)$ when at time t the system was in state $\boldsymbol{x}$ is $\sum_{j=1}^{M} a_j(\boldsymbol{x})dt$. From the law of total probability we can show that:

$$P(\boldsymbol{x}, t+dt \mid \boldsymbol{x_0}, t_0) = P(\boldsymbol{x}, t \mid \boldsymbol{x_0}, t_0)\left(\sum_{j=1}^{M} a_j(\boldsymbol{x})dt\right) + \sum_{j=1}^{M} P(\boldsymbol{x}-\boldsymbol{\nu}_j, t \mid \boldsymbol{x}_0, t_0)a_j(\boldsymbol{x}-\boldsymbol{\nu_j})dt \tag{3.5}$$

Rearranging Equation 3.5, dividing by dt and taking the limit as $dt \to 0$ we obtain the CME:

$$\frac{d}{dt}P(\boldsymbol{x}, t \mid \boldsymbol{x_0}, t_0) = \sum_{j=1}^{M}(P(\boldsymbol{x} - \boldsymbol{\nu}_j, t \mid \boldsymbol{x}_0, t_0)a_j(\boldsymbol{x} - \boldsymbol{\nu_j}) + P(\boldsymbol{x}, t \mid \boldsymbol{x_0}, t_0)a_j(\boldsymbol{x})) \tag{3.6}$$

Equation 3.6 is the chemical master equation (sometimes known as the *forward Kolmogorov equation*). It is a system of k differential equations, where k is the number of possible states of the system. The CME is the most accurate model of stirred reaction systems but unfortunately is it intractable in most situations. For example, a model of a λ phage results in a CME with nearly 10^{70} possible states [Arkin et al., 1998]. It is clear that other models of reaction systems are needed.

3.1.3 The chemical Langevin equation

The Chemical Langevin Equation (CLE) is a slightly less refined approach to reaction kinetics than the CME but numerical solutions are easier to compute. The derivation that we present here was first presented in [Gillespie, 2000]. Assume there exists a time $\tau > 0$ such that the following *Leap condition* is satisfied: $a_j(\boldsymbol{X}(t))$ remains almost constant on the interval $[t, t+\tau)$ for all reactions $1 \leq j \leq M$. The probability that a reaction, R_j occurs in $[t, t+\tau)$ is $a_j(\boldsymbol{X})\tau$. The number of reactions that occur in the interval $[t, t+\tau)$ are Poisson

distributed with parameter $a_j(\boldsymbol{X})\tau$. Assuming that the leap condition holds, the following approximation can be made:

$$\boldsymbol{X}(t+\tau) = \boldsymbol{x} + \sum_{j=1}^{M} \boldsymbol{\nu} P_j(a_j(\boldsymbol{x}), \tau) \tag{3.7}$$

Equation 3.7 is known as the explicit *tau-leaping method* [Gillespie, 2001]. P_j is the probability of a particular reaction occurring. If τ is short enough to make the leap condition hold and long enough that it satisfies $a_j(\boldsymbol{x}) \gg 1$ for all reactions then the Poisson random variable $P_j(a_j(\boldsymbol{x}), \tau)$ can be approximated with a normal random variable with mean and variance $a_j(\boldsymbol{x})\tau$.

$$P_j(a_j(\boldsymbol{x}), \tau) \approx a_j(\boldsymbol{x})\tau + \sqrt{a_j(\boldsymbol{x})\tau} N_j(0,1) \tag{3.8}$$

Equation 3.8 holds when there are sufficiently large populations of reacting chemicals. Substituting Equation 3.8 into Equation 3.7 yields

$$\boldsymbol{X}(t+\tau) = \boldsymbol{x} + \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\boldsymbol{x})\tau + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{a_j(\boldsymbol{x})} \sqrt{\tau} N_j(0,1) \tag{3.9}$$

Equation 3.9 can be approximated by the Euler-Maruyama numerical method to achieve:

$$d\boldsymbol{X}(t) = \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\boldsymbol{X}(t))dt + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{a_j(\boldsymbol{X}(t))} dW_j(t) \tag{3.10}$$

where W_j are Wiener processes. Equation 3.10 is the *Chemical Langevin Equation* in which we have approximated the discrete process of the CME with a continuous process. Unlike the CME whose system of coupled differential equations equalled the number of possible states, the CLE consists of N (where N is equal to the number of species) coupled Ito stochastic differential equations.

3.1.4 The reaction rate equations

The most common approximation to the CME in the biological sciences are the Reaction Rate Equation (RRE). These are useful when the system is within the *thermodynamic limit*, that is the number of molecules and the reaction volume of the system are large ($\boldsymbol{X}_i(t) \to \infty$ and $V \to \infty$). When these limits are applied in Equation 3.9 the stochastic terms impact grows much smaller than the impact of the deterministic component. It is therefore clear that in the thermodynamic limit stochastic terms can be ignored. As we are dealing with large numbers of molecules, we change from a state vector that records number of particles $\boldsymbol{X}(t)$ to one that tracks concentrations $\boldsymbol{Z}(t)$:

$$\boldsymbol{Z}(t) = \begin{pmatrix} Z_1(t) = X_1(t)/VN_A \\ Z_2(t) = X_2(t)/VN_A \\ Z_3(t) = X_3(t)/VN_A \\ \vdots \\ Z_N(t) = X_N(t)/VN_A \end{pmatrix} \tag{3.11}$$

where N_A is Avogadro's number. This concentration vector then satisfies:

$$\frac{d\boldsymbol{Z}(t)}{dt} = \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\boldsymbol{Z}(t)) \tag{3.12}$$

Equation 3.12 is known as the rate equation; the deterministic approximation to the CME. This system of Ordinary Differential Equation (ODE)s is formed from the stoichiometric reaction equations that make up the cellular process. Each reaction equation in the system affects the rate of change of one or more of the chemical species in the system. We combine these rates, both additive and destructive, for each species to produce one ODE per species. The state of the system at any t can be found from $\boldsymbol{Z}(t)$ and as the system is totally deterministic the entire trajectory of $\boldsymbol{Z}$ can be calculated from any state, $\boldsymbol{Z}(t)$.

Despite its position as the most coarse grained approximation to the CME the RRE has enjoyed a great deal of success modelling a wide variety of pro-

cesses in biochemistry and systems biology. Large parts of our understanding of cellular dynamics have come from RRE simulations [Covert et al., 2001; Fell and Cornish-Bowden, 1997; Kauffman et al., 2003]. The RRE has been successfully used to study biochemical oscillations [Tyson et al., 2003] such as the cell cycle [Tyson, 1991] and circadian rhythms [van Zon et al., 2007]. Deterministic modelling by the RRE remains the model of choice for most biological systems [Andrews et al., 2009].

So far in this chapter we have given derivations of the CME and two approximations to this, the CLE and the RRE. We have not yet discussed how these models are solved. In the next chapter we discuss solution techniques, paying particular attention to the solution of ODEs as this is the reaction modelling technique used in this thesis.

3.1.5 Numerical solutions of the chemical master equation

As mentioned above the CME is very difficult to solve for most systems. However progress on numerical methods to solve the CME has been made. One such technique is known as the Finite State Projection (FSP) algorithm which approximates the solution of the CME [Munsky and Khammash, 2006]. The family of exact methods are perhaps the most well known simplifications of the CME. Rather than computing all reaction trajectories of all possible states only one trajectory is simulated. This is achieved by sampling the probability of each reaction S_i occurring and the time that it occurs from a probability distribution consistent with the CME. This approach had generated two stochastic simulation algorithms Stochastic Simulation Algorithm (SSA) to solve the CME; the Direct Method [Gillespie, 1976] and the First Reaction Method [Gillespie, 1977]. These two algorithms are equivalent, both have a time complexity of $O(M)$ where M is the total number of reactions. The Direct Method (also known as Gillespie's algorithm) is more efficient than the First Reaction method, requiring two random numbers per iteration, rather than M. There have been many further refinements to these two exact simulation algorithms including

the Gibson-Bruck algorithm [Gibson and Bruck, 2000] along with many "tau-leaping" methods which advance the time step further than is usual by using an a priori chosen time step - which still satisfies the Leap condition - rather than one from the probability distribution.

3.1.6 Numerical solutions of the chemical Langevin equation

Numerical methods for solving Stochastic Differential Equation (SDE)s are not as numerous as those for solving ODEs. As such, there are only a few numerical methods suitable for solving the CME. The Euler-Maruyama Method (EMM) method discussed in reference to Equation 3.9 and Equation 3.10 is one such method. This technique is an extension of the well known Euler method, from the solution of ordinary differential equations to the solution of stochastic ordinary differential equations. The EMM method is unique in that it is the only method that is so simply extended from the deterministic to the stochastic [Kloeden and Platen, 1992]. We present a brief overview of this technique below. Consider the SDE in Equation 3.13:

$$dX_t = f(X_t)dt + g(X_t)dW_t \tag{3.13}$$

where W_t represents the Wiener process, and $f(x)$, $g(x)$ are defined functions of x. We wish to solve this SDE on the time interval $[0, T]$. The EMM approximation of the true solution X is defined by the following scheme:

$$
\begin{aligned}
\Delta t &= T/L \\
\tau_j &= j\Delta t \\
X(\tau_j) &= X_j \\
j &= 1, 2, 3, ..., L \\
X_j &= X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})(W(\tau_j) - W(\tau_{j-1}))
\end{aligned}
\tag{3.14}
$$

The Milstein method (based upon Markov chains) is another technique for numerically approximating the solution to a SDE [Mil'shtejn, 1975]. As with the EMM we begin with (3.13)

We wish to solve this SDE on the time interval $[0, T]$. The Milstein approximation of the true solution X is defined by the following scheme:

$$
\begin{aligned}
\Delta t &= T/L \\
\tau_j &= j\Delta t \\
X(\tau_j) &= X_j \\
j &= 1, 2, 3, ..., L \\
X_j &= X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})\Delta W + \\
&\quad \frac{1}{2} g(X_{j-1}) g'(X_{j-1})((\Delta W)^2 - \Delta t)
\end{aligned}
\tag{3.15}
$$

where $\Delta W = W(\tau_j) - W(\tau_{j-1})$. The Milstein method is superior in strong order of convergence compared to the EMM, but identical in weak order of convergence [Mackevicius, 2013].

The Runge-Kutta Method (RKM) numerical method for SDEs is an adaptation of the same method for solving ODEs. We begin with the Ito stochastic differential equation:

$$dX_t = f(X_t)dt + g(X_t)dW_t \tag{3.16}$$

where W_t represents the Wiener process, and $f(x)$, $g(x)$ are defined functions of x. We wish to solve this SDE on the time interval $[0, T]$. The Runge-Kutta approximation to the true solution X is defined by the following scheme:

$$\begin{gathered}
\Delta t = T/L \\
\tau_j = j\Delta t \\
X(\tau_j) = X_j \\
j = 1, 2, 3, ..., L \\
X_j = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})\Delta W + \\
\frac{1}{2}\Big(g(\chi_{j-1}) - g(X_{j-1})\Big)\Big((\Delta W)^2 - \Delta\Big)\Delta t^{-1/2}
\end{gathered} \tag{3.17}$$

where $\Delta W = W(\tau_j) - W(\tau_{j-1})$ and $\chi_{j-1} = X_{j-1} + f(X_{j-1})\Delta t + g(X_{j-1})\Delta t^{1/2}$

3.2 Numerical solutions of the reaction rate equation

As previously stated the RRE is seen as the standard for biochemical simulation. The RRE is composed of a system of differential equations and so in order to solve the RRE we must solve the ODEs from which it is composed. Systems that can be modelled by ODE occur frequently in nature, not just in the biological sciences but in disciplines as varied as the physical sciences, engineering, machine learning and economics. Unfortunately solutions to these differential equations can rarely be expressed in closed form. Unfortunately, there is often no way of knowing if an ODE, or other type equation has a closed form solution without trying to find one. On other other occasions, it may be that one exists, but it has not yet been discovered. Therefore we often need to approximate

solutions to these equations with numerical methods. In this section we discuss numerical methods used for integration of ordinary differential equations such as Equation 3.18.

$$\frac{dy}{dx} = f(x, y) \tag{3.18}$$

In simulations performed in this thesis equations such as (3.18) initial values are known - for instance the concentration of a chemical species or its location. These are known as Initial Value Problem (IVP). are We begin with one of the simplest numerical methods used to approximate differential equations, the forward Euler method. Given the initial condition $y(x_0) = y_0$, Euler's method approximates the solution to Equation 3.18 at the subsequent points $x_1, x_2, x_3, ..., x_n$. These points differ by a value h:

$$\begin{aligned} x_1 &= x_0 + h \\ x_2 &= x_1 + h \\ &... \\ x_n &= x_{n-1} + h \end{aligned} \tag{3.19}$$

Further, let y_i denote the approximation to the true solution at the point x_i. Consider the slope of the curve $y'(x_0)$ at the point x_0, y_0. By Equation 3.18:

$$y'(x_0) = f(x_0, y_0) \tag{3.20}$$

Euler's approximation leads to:

$$
\begin{aligned}
y_1 &= y_0 + hf(x_0, y_0) \\
y_2 &= y_1 + hf(x_1, y_1) \\
&... \\
y_{n+1} &= y_n + hf(x_n, y_n)
\end{aligned}
\tag{3.21}
$$

In summary, the numerical method is given by:

$$
y(x+h) = y(x) + hy'(x) \tag{3.22}
$$

The forward Euler method approximates $x + h$ with an error equal to h^2. Another related technique is the backward Euler method; instead of evaluating the function at x the backwards Euler evaluates the function at $x + h$. The backwards Euler method is known as an implicit solution. Implicit schemes are generally considered to be more stable than explicit schemes. This is due to the fact that explicit schemes are conditionally stable; they give accurate answers as long as a specific relationship between timestep length and other quantities is maintained. Timestep length is constrained due to truncation errors building up over time. This often results in longer solving times. Implicit schemes do no suffer the same constraints.

We now review the Runge-Kutta 4th Order Method (RK4) numerical method - the method that we use in our numerical simulation package for solving ordinary differential equations. RKM methods are numerical approximations to differential equations that are more accurate than the Euler approximation, whilst at the same time they are easier to implement than Taylor methods [Atkinson et al., 2011]. First-order initial value problems, such as reaction systems, are usually modelled using either linear multistep or RKM. The "4" represents the fact that a 4^{th} order Taylor series expansion was used in the derivation of this scheme. This formulation of RK4 is taken from [myphysicslab.com, 2015].

Consider the first order differential equation:

$$\frac{dx}{dt} = f(t, x) \tag{3.23}$$

with initial condition $x(0) = x_0$ and at time t_n, $x = (x_n)$. The RK4 takes an x_n, t_n pair and calculates an approximation for x_{n+1} at the next time step, t_{n+1} of size h. We define $h = t_{n+1} - t_n$. x_{n+1} is calculated from a weighted average of f(t,x) at several points within the interval (t_n, t_{n+h}). The value of x_{n+1} is given by:

$$x_{n+1} = x_n + \frac{h}{6}(a + 2b + 2c + d) \tag{3.24}$$

where the coefficients a, b, c, d are given by:

$$\begin{aligned} a &= f(t_n, x_n) \\ b &= f(t_n + \frac{h}{2},\, x_n + \frac{h}{2}a) \\ c &= f(t_n + \frac{h}{2},\, x_n + \frac{h}{2}b) \\ d &= f(t_n + h,\, x_n + hc) \end{aligned}$$

Equation 3.24 is the RK4 for one variable, x. However, most biochemical systems have far more than one reacting chemical. When solving *systems* of differential equations, the variables are stored in vectors.

Suppose that there are m reacting species $x_1, x_2, x_3, ..., x_m$. As stated above, the interactions for each species are combined to produce one differential equation per species, resulting in m ODE:

$$x_1' = f_1(x_1, x_2, x_3, ..., x_m)$$

$$x_2' = f_2(x_1, x_2, x_3, ..., x_m)$$

$$x_3' = f_3(x_1, x_2, x_3, ..., x_m)$$

$$...$$

$$x_m' = f_n(x_1, x_2, x_3, ..., x_m)$$

We collect these species and reaction ODEs together into two vectors. One of species:

$$\boldsymbol{X} = \{x_1, x_2, x_3, ..., x_m\}$$

and a "vector of functions",

$$\boldsymbol{f} = \{f_1, f_2, f_3, ..., f_n\}$$

Now the states used in a multi-variable scheme are $\boldsymbol{X_n}$ and $\boldsymbol{X_{n+1}}$ separated by an interval of h. The $\boldsymbol{X_n}$ vector contains the concentration values of the m species at time step n:

$$\boldsymbol{X}_n = \{x_{1,n}, x_{2,n}, x_{3,n}, ..., x_{m,n}\}$$

$$\boldsymbol{X}_{n+1} = \{x_{1,n+1}, x_{2,n+1}, x_{3,n+1}, ..., x_{m,n+1}\}$$

Now, if the system is in state $\boldsymbol{X_n}$, to advance to state $\boldsymbol{X}_{n+1}$ we calculate:

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \frac{h}{6}(\boldsymbol{a}_n + 2\boldsymbol{b}_n + 2\boldsymbol{c}_n + \boldsymbol{d}_n) \tag{3.25}$$

where the coefficients $\boldsymbol{a}_n, \boldsymbol{b}_n, \boldsymbol{c}_n, \boldsymbol{d}_n$ are calculated from:

$$\boldsymbol{a}_n = \boldsymbol{f}(\boldsymbol{X}_n)$$
$$\boldsymbol{b}_n = \boldsymbol{f}(\boldsymbol{X}_n + \frac{h}{2}\boldsymbol{a}_n)$$
$$\boldsymbol{c}_n = \boldsymbol{f}(\boldsymbol{X}_n + \frac{h}{2}\boldsymbol{b}_n)$$
$$\boldsymbol{d}_n = \boldsymbol{f}(\boldsymbol{X}_n + \frac{h}{2}\boldsymbol{c}_n)$$

RK4 is a popular method of approximating ODEs and, despite enjoying less success on stiff systems, is suitable for all of the systems that we use in this work. We show this to be true by verifying all simulated reaction systems against MATLABs explicit ode45 solver in Chapter 4. Matlab documentation suggests that if ode45 solves the system promptly then the system is not stiff [37]. In addition to this, we verify our results against Mathematicas Wolfram Alpha ODE solver and it is once again demonstrated that a non-stiff solver is suitable for the systems used in this work.

3.3 Mathematical Models of diffusion

It is clear that intracellular distance, organelle positioning, species positioning, and cell shape are all spatial factors capable of influencing a reaction pathway. A model is needed that takes into account these factors. Reactions have been modelled by the techniques described above, but diffusion, an intrinsic physical phenomenon, often remains neglected. Reaction-diffusion models extend reaction systems by adding this transport system - they show how chemical concentrations change under the influence of reaction and diffusion; the net movement of a substance (e.g., an atom, ion or molecule) from a region of high concentration to a region of low concentration. It is important to note that when one of these processes is faster than the other it is said to become *dominated* by the faster process; some reaction-diffusion processes are said to be

reaction-dominated or diffusion-dominated.

There are different ways of modelling reaction-diffusion systems. Often the simplest method of creating a reaction-diffusion model is by adding a diffusive component to an already tested reaction model. SSA approaches can be made spatial by discretising the reacting space and treating a diffusive event, from one spatial quantum to another, as a reaction. The CLE can be modified by adding a term for white Gaussian noise to simulate diffusion [Andrews et al., 2009]. The RRE is subject to perhaps the most recognisable modification; adding a diffusive process modelled with a Partial Differential Equation (PDE). Finally, there exist particle based reaction-diffusion models. These systems represent particles explicitly, with physical laws, such as Molecular Dynamics (MD) or Brownian Dynamics (BD) (or approximations thereof) dictating the evolution of the system. As simulations in this thesis use the RRE there is concentration on PDE based models.

3.3.1 PDE Based Reaction-Diffusion Models

PDE based reaction-diffusion models (that is reaction-diffusion models that use Fick's second law as a transport mechanism) are a suitable tool to simulate reaction-diffusion systems inside cells [Neves and Iyengar, 2009]. PDE reaction-diffusion models allow us to simulate spatial gradients by taking into account the localisation of chemical species and cellular morphology with easily expressed boundary conditions.

As Fick's second law extends the RRE system, comparisons are naturally made between PDE models and ODE models; ODEs describe the change in concentration of a set of species over time, whereas PDE models describe changes over time and space [Eungdamrong and Iyengar, 2004]. PDE models take account of the diffusive processes that occur within a system that results in a change of concentration. PDE systems are more challenging to implement with very few large scale projects available for use [Neves and Iyengar, 2009]. VCell [Loew and Schaff, 2001] is one such simulation package.

Mathematically, we describe a PDE reaction diffusion system as:

$$\frac{\partial u}{\partial t} = D\nabla^2 u + R(u) \tag{3.26}$$

or alternatively:

$$\frac{\partial u}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + R(u) \tag{3.27}$$

Where u is the concentration, D is the diffusion coefficient and R accounts for a system of reactions described by a set of ODEs. As mentioned earlier we use Neumann boundary conditions. Solving this system gives a discretised grid of concentration values that varies over time. To do this, we need to solve both PDEs ($\nabla^2 u$) and ODEs ($R(u)$).

In real cellular systems, there can be many different chemical species. As all species must be allowed to diffuse, this equation has to be solved for each species. This can become very computationally expensive.

3.3.2 Solving PDE Reaction-Diffusion Models

Equation 3.26 has both an ODE component and a PDE component. As we have discussed solution methods for ODEs we now describe methods for numerically solving PDEs. Hyperbolic PDEs, such as the diffusion equation, can be solved in many different ways, two of the most common are the explicit finite difference method and the implicit Crank-Nicholson method. In three dimensions, the explicit method is bound by the stability condition: $\frac{dt}{dx^2} < \frac{1}{6}$; that is, the size of the timestep divided by the square of the smallest measurable distance must be less than $\frac{1}{6}$.

The choice of dx is important as it dictates the level of detail that can be extracted from our simulation. Just as dt is the smallest unit of measurable time, dx is the smallest unit of measurable space. It can be thought of as the

maximum spatial resolution. If one wanted to model a cellular process at the level of an organelle, then dx would need to be at least as small as the organelle in question. A smaller dx gives a more detailed simulation. The explicit condition would then require that the timestep be made even smaller, which increases the time required for the simulation. A longer simulation would be possible but require more time or a more powerful machine.

The alternative to using an explicit scheme is an implicit scheme, where the reaction diffusion system is solved as a set of simultaneous equations by a linear solver. Implicit methods do not have to observe the same stability condition however, implementing boundary conditions (such as complex cell and organelle shapes) in an implicit regime is much more difficult.

We now derive the explicit scheme (from [Smith, 1985]) used for the majority of experiments in this thesis: If a function describing concentration, U, and its derivatives are single-valued, finite and continuous functions of position (x), then a Taylor expansion of U yields:

$$U(x+h) = U(x) + hU'(x) + \frac{1}{2}h^2U''(x) + \frac{1}{6}h^3U'''(x) + ... \tag{3.28}$$

and

$$U(x-h) = U(x) - hU'(x) + \frac{1}{2}h^2U''(x) - \frac{1}{6}h^3U'''(x) + ... \tag{3.29}$$

Summing Equation 3.28 and Equation 3.29 gives:

$$U(x+h) + U(x-h) = 2U(x) + h^2U''(x) + O(h^4) \tag{3.30}$$

$O(h^4)$ represents terms in the expansion of powers of h of 4 and greater, which we assume have a negligible impact upon the solution. Rearranging Equa-

tion 3.30 gives us:

$$U''(x) = \frac{d^2U}{dx^2} \approx \frac{1}{h^2}\Big\{U(x+h) - 2U(x) + U(x-h)\Big\} \tag{3.31}$$

Subtracting (3.29) from (3.28) and rearranging in a similar way gives:

$$U'(x) = \frac{d^2}{dx} \approx \frac{1}{2h}\Big\{U(x+h) - U(x-h)\Big\} \tag{3.32}$$

Equation 3.32 is an approximation of the tangent at the point x and is known as the *central-difference* approximation. We can also think about the gradient at x in two other ways; between $U(x+h)$ and $U(x)$ known as the *forward-difference* approximation:

$$U'(x) = \frac{d}{dx} \approx \frac{1}{h}\Big\{U(x+h) - U(x)\Big\} \tag{3.33}$$

and between $U(x)$ and $U(x-h)$ known as the *backwards-difference* approximation:

$$U'(x) = \frac{d}{dx} \approx \frac{1}{h}\Big\{U(x) - U(x-h)\Big\} \tag{3.34}$$

Assuming that U is a function of both x and t and discretising the solution space, $\delta x = h$, $x_i = ix$, $i = 0, \pm1, \pm2, ...$ and $\delta t = k$, $t_j = jt$, $j = 0, \pm1, \pm2, ...$ then Equation 3.31 can be rewritten in this new notation as:

$$\frac{d^2U_x^t}{dx^2} \approx \frac{U_{x+1}^t - 2U_x^t + U_{x-1}^t}{h^2} \tag{3.35}$$

If the environment being simulated is a 3D environment then we can write down similar equations to Equation 3.35 (assuming the discretisation is the same in all dimensions) for the y and z dimensions:

$$\frac{d^2U_y^t}{dy^2} \approx \frac{U_{y+1}^t - 2U_y^t + U_{y-1}^t}{h^2} \tag{3.36}$$

$$\frac{d^2U_z^t}{dz^2} \approx \frac{U_{z+1}^t - 2U_z^t + U_{z-1}^t}{h^2} \tag{3.37}$$

We now have finite difference schemes for $\frac{d^2U}{dx^2}$ $\frac{d^2U}{dy^2}$ $\frac{d^2U}{dz^2}$ $\frac{d^2U}{dt^2}$. This means we are in a position to form a finite difference approximation for the diffusion terms of the reaction-diffusion equation. The diffusion terms are:

$$\frac{\partial u}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \tag{3.38}$$

We can substitute Equation 3.35, Equation 3.36 Equation 3.37 into Equation 3.38 to obtain a finite difference approximation to Equation 3.38. This results in:

$$\begin{aligned} U_{x,y,z}^{t+k} = r[U_{x+h,y,z}^t + U_{x-h,y,z}^t] + r'[U_{x,y+h,z}^t + U_{x,y-h,z}^t] + \\ r''[U_{x,y,z+h}^t + U_{x,y,z-h}^t] + (1 - 2r - 2r' - 2r'')U_{x,y,z}^t \end{aligned} \tag{3.39}$$

The finite difference approximation of the diffusion equation, where $r = \frac{k}{h_x^2}$, $r' = \frac{k}{u_y^2}$ and $r'' = \frac{k}{u_z^2}$. Equation 3.39 is an explicit difference scheme and is unstable when the condition $\frac{dt}{dx^2} < \frac{1}{6}$ is not met. The Crank-Nicholson method [Juncosa and Young, 1957] is an implicit scheme which is always stable regardless of the time step size. Crank-Nicholson is formulated by taking the mean of the j and $j+1$ time steps. In one dimension:

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} \tag{3.40}$$

substituting finite difference approximations into Equation 3.40:

$$\frac{U_x^{t+1} - U_x^t}{k} = \frac{1}{2}\left[\frac{U_{x+1}^{t+1} - 2U_x^{t+1} + U_{x-1}^{t+1}}{h^2} + \frac{U_{x+1}^t - 2U_x^t + U_{x-1}^t}{h^2}\right] \tag{3.41}$$

which can be rearranged into

$$-rU_{x-1}^{t+1} + (2+2r)U_x^{t+1} - rU_{x+1}^{t+1} = rU_{x-1}^{t} + (2+2r)U_x^{t} - rU_{x+1}^{t} \qquad (3.42)$$

Providing boundary and initial conditions are known, a system of Equation 3.42 can be constructed, one for each point in the simulation mesh. This system is usually formulated as a matrix and solved as an $A\boldsymbol{x} + \boldsymbol{b} = 0$ system using a linear solving technique such as elimination of variables or Cramer's rule.

3.4 Summary

We have discussed a wide variety of formulations and solutions of reaction-diffusion models. In the simulations presented in this thesis, the reaction-diffusion equations are solved by dividing the equation into PDE and ODE components, solving them separately and summing them in a technique known as operator splitting [Knio et al., 1999]. In this thesis our simulation package uses the explicit finite difference method (Equation 3.39) to numerically solve the diffusion equation and the RK4 algorithm to numerically solve reaction systems. Despite the explicit scheme being slower than implicit scheme it is no less accurate (provided the timestep size bounding condition is followed). In addition, memory limitations on some of General Purpose Graphics Processing Unit Computing (GPGPU)s available meant that for larger systems, the implicit method was infeasible. For this reason the explicit scheme was chosen over the explicit scheme. The simulation infrastructure is discussed in Chapter 4.

CHAPTER 4

ReDi-Cell

4.1 Introduction

In this chapter, we introduce Reaction-Diffusion Cell (ReDi-Cell), the simulation package that we have developed for this thesis. All results in this thesis are generated from simulations using this environment. We begin by discussing two reaction systems that make suitable abstractions for cellular pathways. The architecture of ReDi-Cell is then discussed; choice of numerical methods, visualisation capabilities and parallelism provided by GPU support are detailed. Finally we demonstrate the accuracy of our newly developed simulation by testing it against Virtual Cell Simulation Package (VCell), an advanced cell simulation package. ReDi-Cell is found to produce very similar results to VCell.

4.2 Simulated abstractions of cellular pathways

In this section, we give details of the two reaction-systems, simulated in the ReDi-Cell environment, that are used as abstractions for *in vivo* pathways. The two reaction systems simulated are the Lotka-Volterra (LV) system and the Bi-Molecular Reaction (BMR) system. Abstractions can be used when the details of the reaction system are still controversial, many models are based on hypothetical pattern-formation systems [Holmes and Edelstein-Keshet, 2012].

4.2.1 Representative abstractions for cellular pathways

Despite the accomplishments of technologies such as fluorescence photo-bleaching, the parameters for many chemicals in many biological pathways are unknown. This work is not concerned with the behaviour of a *specific* reaction system, but

rather the effect of cell geometry on a *general* reaction system. At a fundamental level, all cellular processes consist of coupled chemical reactions. Biological pathways are no exception to this; therefore we are able to model any cellular pathway (at an abstract level) with a general reaction system coupled with diffusion. Accordingly we simulate *abstracted* cellular pathways and in so doing we avoid the need for parameters, while at the same time still representing the way in which geometry can affect cellular pathways.

In vivo biological pathways are often highly complicated, involving many reactions and reactants. *In silico* representations of these pathways are often simplified, abstracted versions of the real thing. For example, the NF-kB pathway has been abstracted in several studies [Krishna et al., 2006; Terry, 2014]. When we seek to answer new questions or explain new behaviours, simpler, abstracted systems can be advantageous [Krishna et al., 2006].

Two reaction systems are selected as the abstracted cellular pathways; LV [Lotka, 1925; Volterra, 1927] and BMR. By using dissimilar reaction systems general conclusions as to the impact of geometry on cellular pathways can be drawn, that are not due to artefacts of one of the abstractions.

By selecting a proven abstraction of an oscillatory reaction system, and a simpler system, that forms the basis for many more complex systems, a good representation of reaction systems found in nature can be obtained. However, it is important to note that while the two reaction systems selected can approximate a wide variety of reaction systems, there are some that they cannot.

4.2.2 Lotka-Volterra

The LV system was proposed as a model for both autocatalytic reactions, by Lotka, and for population dynamics, by Volterra. Over time it became a ubiquitous example of a predator-prey system [McLaughlin and Roughgarden, 1991] with its qualities and variations explored extensively in the literature.

LV systems are a representative of a particular class of biochemical reaction known as an *oscillating* reaction - it has been used as an abstraction for this type

of reaction system in the literature [Andrews et al., 2009]. Many systems within the cell are understood to be controlled by oscillating reaction systems [Vîlcu and Bala, 2004], with circadian rhythms [Noyes and Field, 1974], insulin and hormone secretions being well known examples. At shorter time scales, there are several signalling mechanisms [Novák and Tyson, 2008] that qualify as oscillating systems. Specific examples include mitosis control in the slime mold [Tyson and Kauffman, 1975] and glycolysis within yeast cells [Hess et al., 1973]. The NF-kB pathway (discussed above) also displays oscillatory behaviour[Foreman et al., 2004; Friedrichsen et al., 2006; Ihekwaba et al., 2004]. There are several theorems that describe the necessary conditions for oscillatory reactions to occur within bi-molecular reaction systems [Heinrich and Schuster, 2012]. Here we will list two that describe the conditions completely. The first condition is that there has to be more than one chemical species participating in the system. Rates of change for the species will be unique solutions, and so cannot have opposing signs [Heinrich and Schuster, 2012]. The second condition is that chemical systems, containing only two species, that display only uni-molecular or bi-molecular reactions cannot oscillate [Heinrich and Schuster, 2012].
The LV) system is mathematically described as:

$$\frac{dx}{dt} = x(\alpha - \beta y) \tag{4.1}$$

$$\frac{dy}{dt} = -y(\gamma - \delta x) \tag{4.2}$$

Where x is the number of prey, y is the number of predators, $\frac{dx}{dt}$ and $\frac{dy}{dt}$ are the rates of reproduction of the populations. α, β, γ, δ are parameters describing the reproduction and death rates of the two species.

LV systems have previously been coupled with diffusion (Lotka-Volterra-Diffusion (LVD)) [mod, 2014; Hastings, 1978]. Research into LVD)systems is split between investigating mathematical properties of the system, such as solubility or stability analysis, and computational work such as ecological migration.

4.2.3 Bi-Molecular Reaction System

The simplest system that we use is the BMR reaction system and it one of the most common types of reaction system within the cell [Chen et al., 2010]. Known as a second order *synthesis* reaction it is one of the basic types of reaction system along with *decomposition* and *replacement*. Despite its simplicity, it is still able to describe many interactions in real cellular systems. Examples of bimolecular reaction systems include glycolysis, lypolysis, the binding of a transcription factor to Deoxyribonucleic acid (DNA)) the formation of a heterodimer through the binding of two different proteins, and the conversion of a substrate into a complex [Chen et al., 2010].

The BMR system is described as:

$$\frac{da}{dt} = -kab \tag{4.3}$$

$$\frac{db}{dt} = -kab \tag{4.4}$$

$$\frac{dc}{dt} = kab \tag{4.5}$$

a, b, c are the concentration of species a,b and c, and k is the rate constant of the reactions.

4.3 ReDi-Cell

In chapter 2, we saw that the cells display heterogeneous internal and external geometries, and that accurate representation of cell structure is crucial to simulation. Chapter 3 demonstrates that whilst computationally expensive, reaction-diffusion systems are suited to modelling cellular pathways in complex geometries. Finally, in the previous section, we gave evidence that, if we are interested in the effect of cell geometry on a *general* reaction system, then complex

pathways can be replaced with abstractions, such as the LVD system.

These conclusions suggest that cellular pathway simulations should be capable of representing abstractions of reaction-diffusion systems inside complex cell geometry in a performant manner. In this section, we present ReDi-Cell, a novel simulation package that fulfils these requirements.

4.3.1 Description

ReDi-Cell is a 3D high performance General Purpose Graphics Processing Unit Computing (GPGPU) reaction-diffusion simulation of arbitrary cellular geometry. It is the first GPGPU PDE-ODE cell simulation infrastructure specifically designed for investigating the effect of realistic cell shapes and internal components. ReDi-Cell is designed to take cellular morphology information, store it as voxels, and numerically solve a reaction-diffusion system representing physiological processes. Each voxel represents a quantum on a regularly spaced, three-dimensional grid. Results are displayed using Visualise It (VisIt) which is discussed later.

Systems incorporating Partial Differential Equation (PDE)s are more challenging to implement, with few large scale projects available for use [Neves and Iyengar, 2009]. VCell [Loew and Schaff, 2001] is one such simulation package. VCell includes a variety of simulation techniques not limited to PDE simulation and has been used in a variety of published models.

ReDi-Cell, as a high performance GPGPU solver tailor made for PDE-ODE systems, scales to far more detailed problems over more time steps than are feasible with the available VCell interface. The user may dedicate as many GPUs as they have available to the task; spatial decomposition occurs over all available GPUs.

4.3.2 Realistic cell shapes

Bacterial cells appear in nature in a variety of shapes, as shown in Figure 4.1. ReDi-Cell simulates these bacterial shapes by discretising arbitrary cellular mor-

phology, described as 3D models, into voxels. Figure 4.2 demonstrates how geometric primitives have been used to create 3D models to approximate three of the common cell shapes in Figure 4.1. When shapes cannot be described exactly by a single primitive, composites are formed. For example, *Bacilli* are best approximated by the mathematical definition of a capsule; they can be built as a composite of a cylinder and two hemispheres.

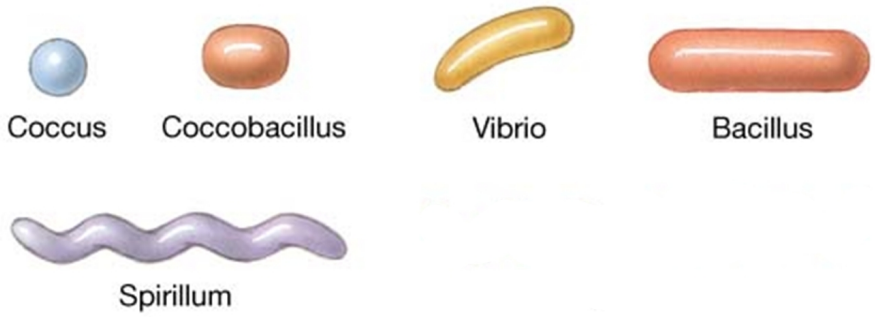
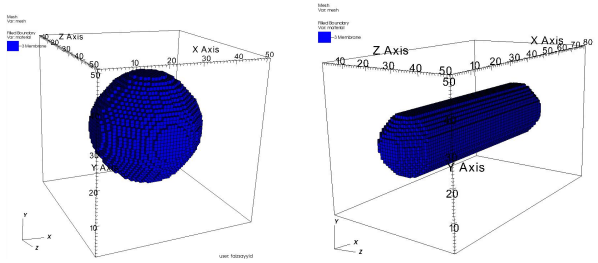
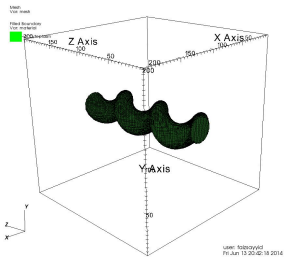


Figure 4.1: Common bacterial shapes reproduced from http://microbeonline.com/characteristics-shape-of-pathogenic-bacteria



(a) *The Coccus* (b) *The Bacillus*



(c) *The Spirillum*

Figure 4.2: Approximations to cell shapes in Figure 4.1 using the VisIt software toolkit. For details of VisIt see text

4.3.3 Cellular Sub-Components

ReDi-Cell allows internal cell components, such as organelles, to be represented. Each component's permeability can be altered by changing the diffusion parameter of that location in the cell. By changing the rate of diffusion at that point in the simulation, organelles can support various levels of permeability. This allows ReDi-Cell components to mimic their biological analogues in the cell, with different reagents having different responses to different cell materials. For instance, some reagents are confined to the plasma membrane, whereas others might be able to pass through it. In Figure 4.5 we see a chemical diffusing in a crowded cellular environment with cellular sub-components. The chemical, incapable of diffusing through the impermeable membrane of the cell nucleus is forced to diffuse around the internal geometry that represents an organelle. Figure 4.3 shows an "onion-peel" decomposition of an example simulation environment. This illustrates the way ReDi-Cell components are capable of capturing the internal structure of the cell. Figure 4.4 shows an example of a 2D

slice through the same ReDi-Cell environment. Figure 4.4 includes grid lines to show the way in which space has been discretised. The level of discretisation is configurable.

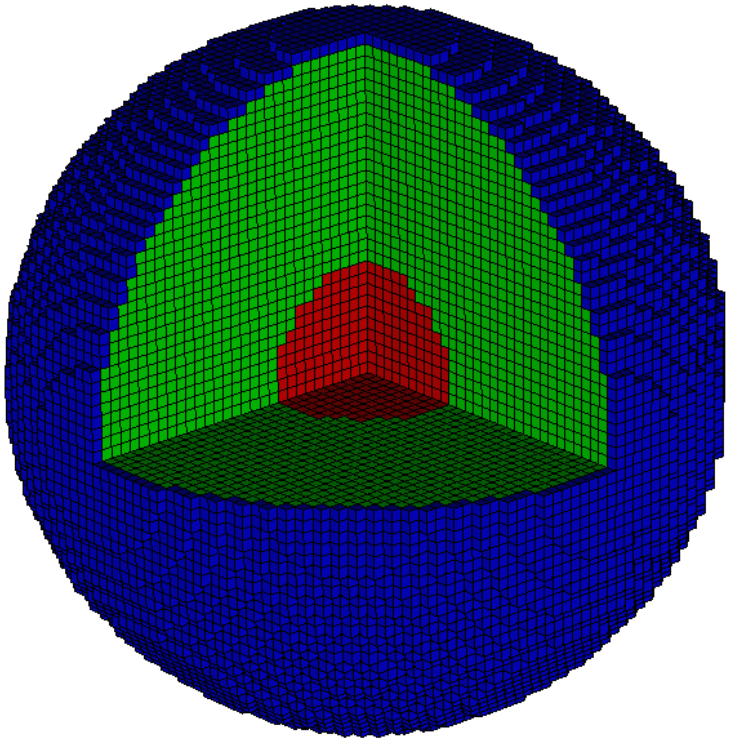


Figure 4.3: Cutaway of a ReDi-Cell model *Coccus* showing the membrane (blue), cytoplasm (green) and nucleus (red) produced using VisIt. Discussion of VisIt follows in the text. The different colours represent different biological material types which have different simulation properties. These simulation properties allow for differing rates of reaction and diffusion in those areas of the cell.

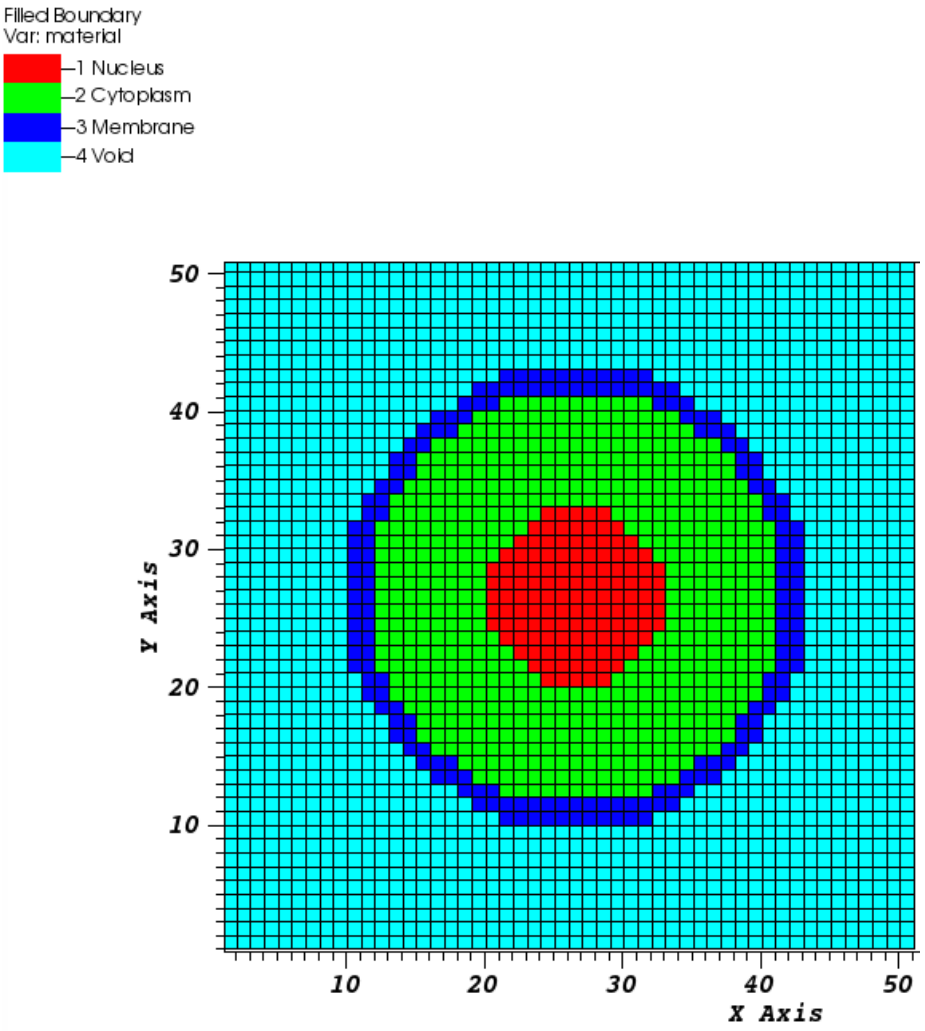


Figure 4.4: 2D slice through the centre of the ReDi-Cell model of a *Coccus* depicted in Figure 4.3. This figure illustrates the way in which VisIt can change perspective to show specific regions of interest. 2D slice results are used extensively in this work.

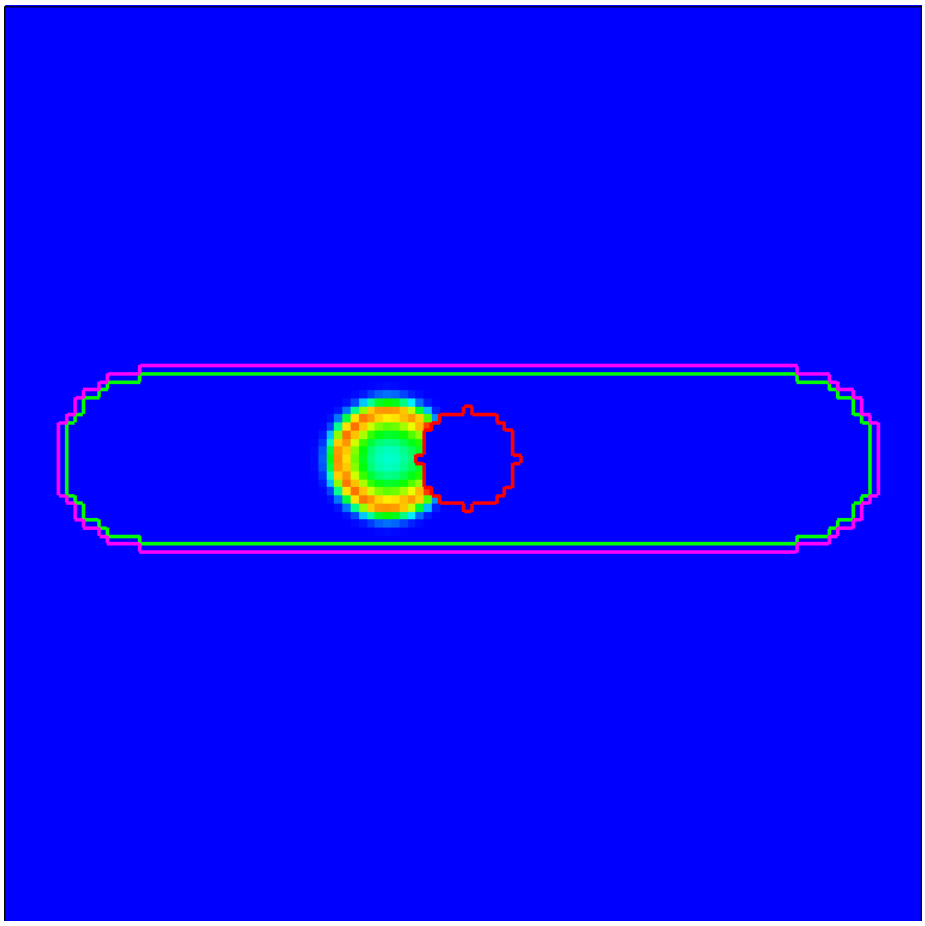


Figure 4.5: 2D slice, produced using VisIt, through the centre of a ReDi-Cellmodel of a *Bacillus*, with a model chemical diffusing around a central organelle. There is no reaction present in this simulation. This image shows how different materials can have different simulation properties and how VisIt can render materials in different ways. Instead of the solid material representations found in Figure 4.4 and Figure 4.3 only the boundaries of materials are drawn allowing both material type *and* chemical concentration to be rendered in the same image. The boundaries of the different biological materials are represented by the different coloured lines. Red represents the boundary of a central organelle, light green line the internal boundary of the cell and purple the external boundary of the cell.The boundaries in the image are overlayed on to a heatmap of chemical concentration, this is the reason that most of the image is dark blue (representing 0 concentration). The multi-coloured circle to the left of the central organelle shows the only non-zero concentration portion of the heatmap; a reacting and diffusing circle of chemical concentration that has not reached the internal boundary. At $t = 0$, this chemical was a ring of concentration. In this case, the material enclosed by the red boundary does not permit reaction or diffusion to occur and the chemical is halted at the membrane. The membrane permits both reaction and diffusion but the rate of diffusion is halved. Beyond the purple line representing the external boundary of the membrane, only diffusion processes are permitted.

It is unlikely that all sub-cellular structures will have homogeneous morphology; organelles are of many different shapes and sizes. ReDi-Cell allows internal cell component shapes to be represented as accurately as the external cell shape. We see an example of this capability in Figure 4.6, which shows spherical and cylindrical shaped organelles inside a *Spirillum*.

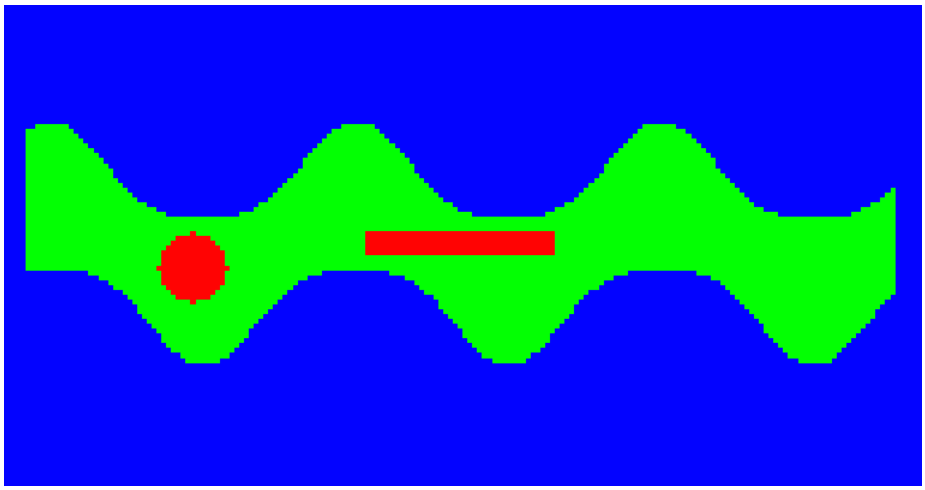


Figure 4.6: 2D slice through the centre of a ReDi-Cell model *Spirillum* produced using VisIt. This figure shows only geometry and no reactions. The boundaries in this image are solid as in Figure 4.3 and Figure 4.4 and so no concentrations are visible. This figure demonstrates different shaped organelles inside the model cell. Green represents the cytoplasm in which chemicals are permitted to react and diffuse. Red shapes represent organelles through which chemicals cannot diffuse or react inside.

4.3.4 ReDi-Cell Architecture

Simulation execution in ReDi-Cell is divided into time steps. A single time step consists of a diffusion process and a reaction process. First, the Ordinary Differential Equation (ODE)s representing the reactions are solved. This process occurs over each part of the cell in which diffusion is allowed to take place. Then the 3D PDE diffusion equation is solved for each reagent, again in the appropriate parts of the cell. This process is known as operator splitting. This technique works by separating the original equation into its constituent terms over a single timestep, computing the solutions to each term, and then combin-

ing the individual solutions to form a solution to the original equation. This technique is commonly encountered when solving convection-diffusion PDEs, where convection and diffusion are solved separately [Glowinski et al., 2017].

There are many different ways of solving PDEs and ODEs, as discussed in Chapter 3. We choose to solve PDEs using an explicit Finite Difference Methods (FDM) algorithm [Smith, 1985] and PDEs using a Runge-Kutta 4th Order Method (RK4) algorithm [Press et al., 2002] using the justifications in the summary of Chapter 3. The FDM algorithm is a common method of simulating reaction-diffusion and is used in many modern studies of such systems [Holmes and Edelstein-Keshet, 2012]. RK4 is an equally popular method of solving ODEs and despite enjoying less success on stiff systems is suitable for all of the systems that we use in this work. We show this to be true by verifying all simulated reaction systems against MATLAB's explicit ode45 solver.

ReDi-Cell uses VisIt for visualisation, a popular open source tool for visualizing and analyzing data [Childs et al., 2012]. VisIt is capable of displaying 2D slices of 3D shapes. In Figure 4.4 and Figure 4.6 we see examples of such cutaway cells.

4.3.5 Implementation

ReDi-Cell has two different implementations of the reaction-diffusion process, a parallel algorithm executed on the GPU and a serial algorithm executed on the CPU for when a compatible GPU is not available.

Reaction-diffusion simulation methods (including the Explicit FDM and RK4 algorithms implemented in ReDi-Cell) are extremely resource intensive. High performance computing architectures, such as the GPGPU targeted Nvidia Compute Unified Device Architecture (CUDA) *www.nvidia.com/cuda*, enable the parallelisation of simulations, allowing faster run times and more detailed models. ReDi-Cell includes a CUDA GPU implementation of the Explicit FDM-RK4 reaction-diffusion method. In the GPGPU algorithm, many reaction or diffusion events are run at the same time, instead of one after the other, as is

the case in the CPU algorithm. In a single time step, all reaction events are run after all diffusion events have taken place. During the simulation, data is written to the disk for offline analysis. Transferring data to and from the GPU is expensive and thus writing to disk every time step can degrade performance. Therefore we compromise, transferring data and writing to the disk at configurable intervals.

The execution of time steps is "batched" together in configurable intervals. Execution is paused between batches whilst results are written to the disk. The larger the batch size the faster the simulation will run. Time step batch sizes of 100 steps are commonly used, which provides a good trade-off between speed and output granularity.

The GPU algorithm has two stages: The kernel launcher, and the kernel itself. The kernel launcher runs on the CPU, it divides the simulation into batches and writes output to the disk. The kernel runs on the GPU and is responsible for executing the reaction process in parallel and then the diffusion processes in parallel across all of the GPU cores.

We also include a CPU algorithm. A single time step consists of a serial implementation of the explicit FDM-RK4 reaction-diffusion method. After each time step, the state of the system is written to disk.

In terms of speed, we find that ReDi-Cell can perform approximately 24x faster than VCell when running a BMR simulation inside a *coccus*. ReDi-Cell benchmark simulations were performed on an NVIDIA Tesla K20.

4.3.6 ReDi-Cell Measurements

ReDi-Cell outputs concentration timeseries and 3D cell visualisations. In the results in this thesis we talk about two related measurements; concentration/reaction trajectory and concentration wavefront shape. We define the reaction/concentration trajectory as the shape of the timeseries of concentration. We define the wavefront as the surface that has, on one side only zero concentration simulation subvolumes and on the other side only non-zero concentration simulation sub-

volumes.

4.4 Experiments

In this section, we detail three sets of experiments that we run with ReDi-Cell. The first is focused on validating ReDi-Cell's implementation against the state of the art, VCell. The validation experiments use two different prototypes: diffusion only and $A + B \rightarrow C$. A summary of the reaction systems is shown in Table 4.1. In the second set, we investigate the behaviour of abstracted cellular pathways in different cell shapes. In the third set, we investigate the impact of clutter on the same pathways. For a discussion of the choice of diffusion coefficients see appendix 1.

System	Reaction Equation	Initial Conditions
Diffusion		A:10
$A + B \rightarrow C$	$\frac{dA}{dt} = -k * A * B$ $\frac{dB}{dt} = -k * A * B$ $\frac{dC}{dt} = k * A * B$	A:10 B:10 C:0

Table 4.1: Description of validation experiments. The parameters used in the experiments are shown in the captions of the figures depicting experimental results. Initial positioning described in results. Concentration is measured in arbitrary units.

Experiments take place in a variety of cluttered and uncluttered environments. The results take the form of concentration heat maps and concentration against time graphs.

Graphs of concentration against time are used when studying the dynamics of a reaction in a single volume. Heatmaps are used when measuring the concentration in many volumes at once, at a single point in time. We exploit this property to measure the effect of space on the distribution of concentrations.

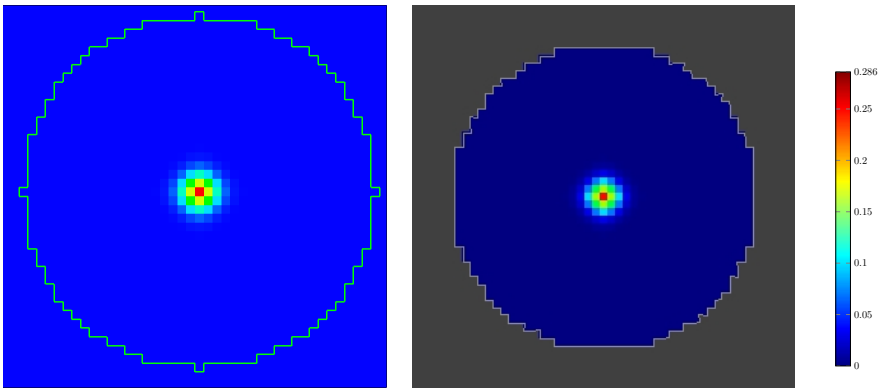
Table 4.2: Cell geometry used in these experiments

Cell	Geometry
Coccus	Diameter: $50\mu m$

Many heatmaps recorded at sequential times can be stitched together into an animation to show the same information as a concentration graph, but for multiple volumes at once. Both concentration against time graphs and heatmaps are common in cell simulation, with VCell (against which we validate our results) using these measurements.

4.4.1 Uncluttered Diffusion

This set of experiments investigates diffusion-only behaviour. The diffusion-only system from Table 4.1 is run in an uncluttered setting in ReDi-Cell. Results from this experiment are shown in Figure 4.7(a) and Figure 4.8. This same system was then run again in an identical uncluttered setting in VCell. Results from this experiment are shown in Figure 4.7(b) and Figure 4.8. These figures show that the results of these experiments are in good agreement.



(a) Heat map of species A concentration in ReDi-Cell

(b) Heat map of species A concentration in VCell

Figure 4.7: Verification Experiment 1 - uncluttered diffusion. Heat maps of species A after 3000 time steps of diffusion. The experiment is run in both ReDi-Cell (Figure 4.7(a)) and VCell (Figure 4.7(b)). It can be seen that the heat maps are identical validating the accuracy of ReDi-Cell's diffusion computation. Colour bar describes concentration in both results. $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

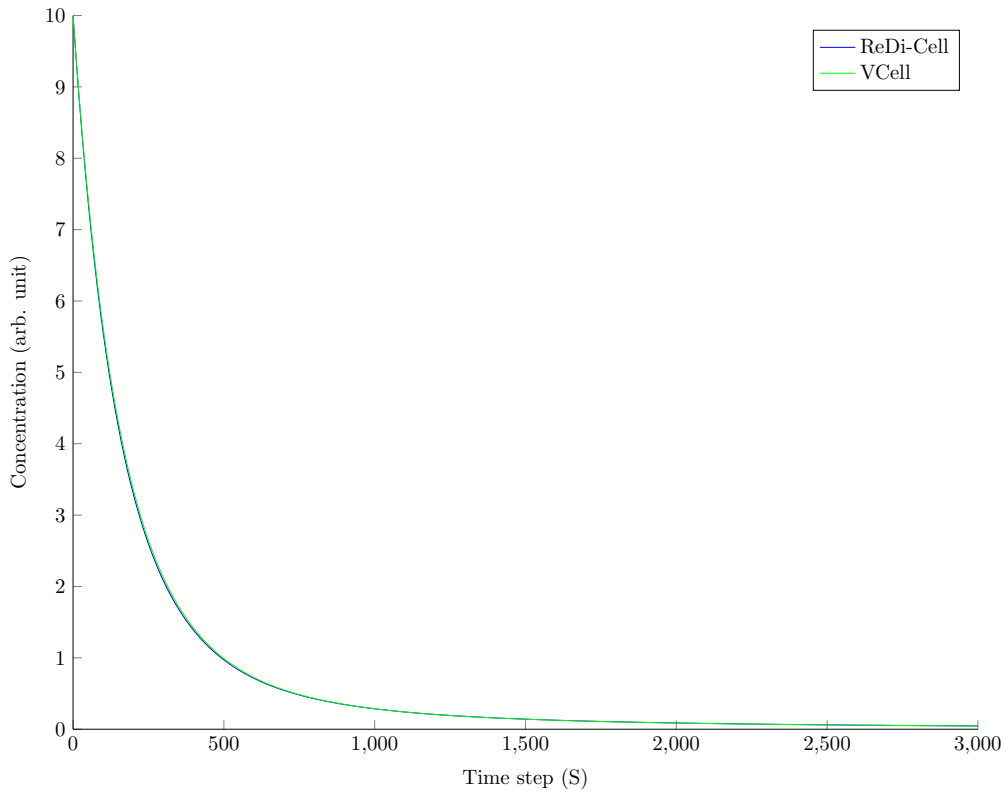
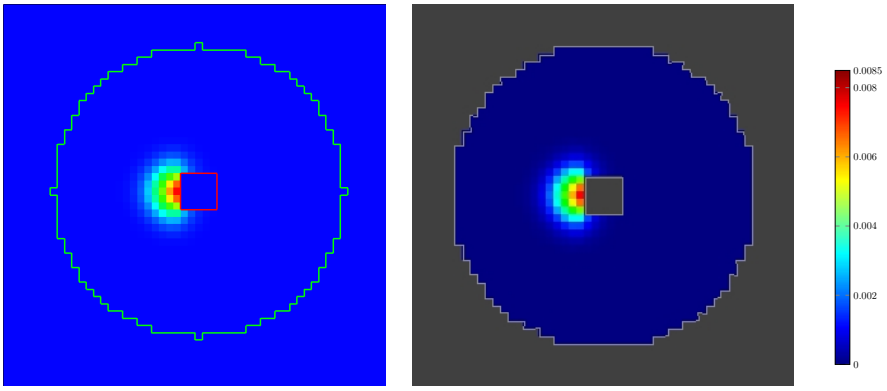


Figure 4.8: Experiment 1 - uncluttered diffusion. Species (A) concentration in ReDi-Cell and VCell measured in a single sub-volume over 3000 time steps. Concentration is measured in the sub-volume that the chemical was initially confined to. The lines that represent the concentration in ReDi-Cell and VCell are on top of each other validating the accuracy of ReDi-Cell's cluttered reaction-diffusion computation. $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

4.4.2 Cluttered Diffusion

This set of experiments examines how ReDi-Cell models the interaction of chemicals with the impermeable boundaries that model the cell boundary and organelles. The diffusion system from Table 4.1 was run in a cluttered environment, consisting of a single impermeable cube. The species are initially located next to the cube. The experiment is first run in ReDi-Cell. Results from this experiment are shown in Figure 4.9(a) and Figure 4.10. The same system was run in VCell. VCell's results are shown in Figure 4.9(b) and Figure 4.10. Once again, the results from ReDi-Cell and VCell are consistent with each other.



(a) Heat map of species concentration in ReDi-Cell

(b) Heatmap of species concentration in VCell

Figure 4.9: Verification Experiment 2 - cluttered diffusion. Heat maps of species A after 3000 time steps of diffusion. The experiment is run in both ReDi-Cell (Figure 4.9(a)) and VCell (Figure 4.9(b)). It can be seen that the heat maps are identical, validating the accuracy of ReDi-Cell's cluttered diffusion computation. $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

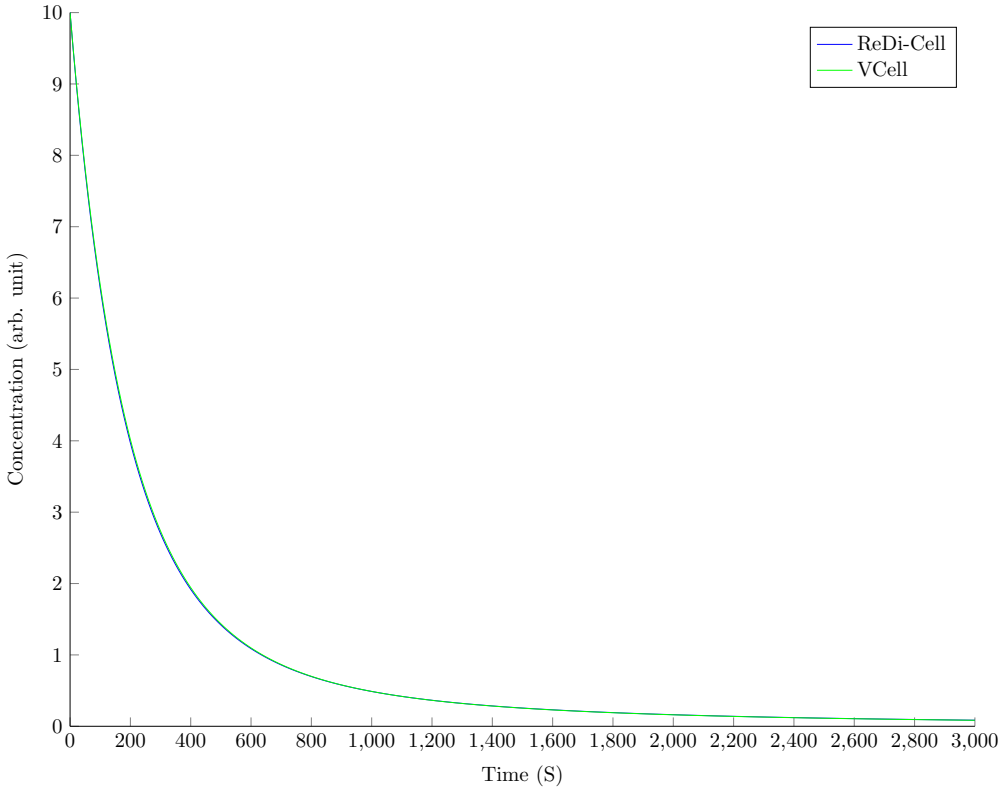
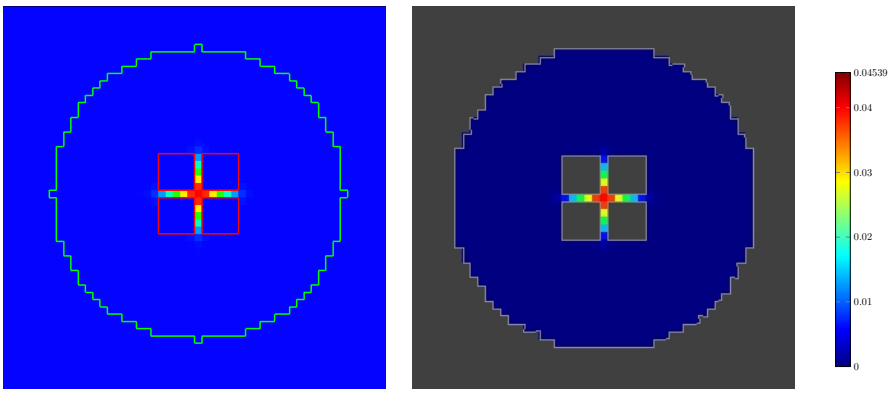


Figure 4.10: Verification Experiment 2 - cluttered diffusion. Graphs of species (A) concentration in ReDi-Cell and VCell measured in a single sub-volume over 3000 time steps. Concentration is measured in the sub-volume that the chemical was initially confined to. The lines that represent the concentration in ReDi-Cell and VCell are on top of each other validating the accuracy of ReDi-Cell's cluttered reaction-diffusion computation. $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

4.4.3 Cluttered $A + B \rightarrow C$

In this set of experiments, we investigate ReDi-Cell's behaviour when simulating reaction-diffusion next to impermeable membranes. The $A + B \rightarrow C$ reaction-diffusion system from Table 4.1 is run in a cluttered environment. This clutter takes the form of four impermeable cubes, separated by narrow channels. The species are initially located at the centre of cell and at the point of intersection of the channels. The experiment is first run in ReDi-Cell, the results of this experiment are shown in Figure 4.11(a) and Figure 4.12. This experiment is run again in an identical environment in VCell. The results of this experiment are shown in Figure 4.11(b) and Figure 4.12. ReDi-Cell's results are shown to be consistent with VCell's.



(a) Heatmap of species C concentration in ReDi-Cell

(b) Heatmap of species C concentration in VCell

Figure 4.11: Verification Experiment 3. Cluttered reaction-diffusion of the $A + B \rightarrow C$ system. The experiment is run in both ReDi-Cell (Figure 4.11(a)) and VCell (Figure 4.11(b)). It can be seen that the heat maps are identical, validating the accuracy of ReDi-Cell's cluttered reaction-diffusion computation. $k = 1.4^{12}s^{-1}$, $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

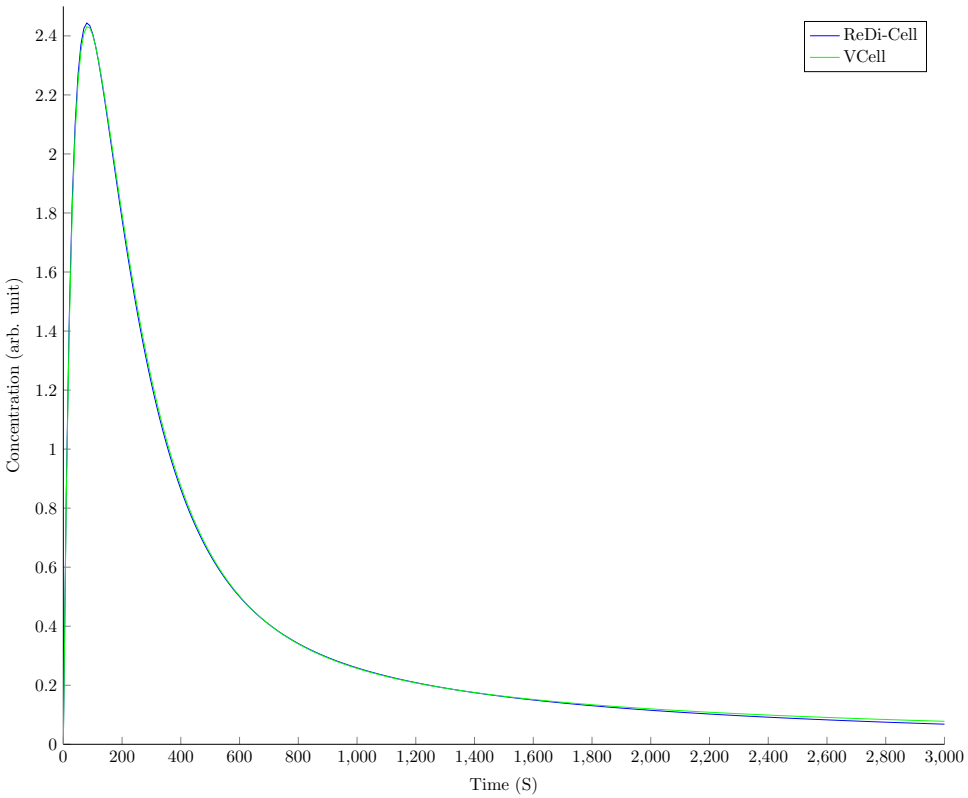


Figure 4.12: Verification Experiment 3 - Cluttered reaction-diffusion of the $A + B \rightarrow C$ system. Graphs of species C concentration in ReDi-Cell and VCell measured in a single sub-volume over 3000 time steps. Concentration is measured in the sub-volume that the chemical was initially confined to. The lines that represent the concentration in ReDi-Cell and VCell are on top of each other, validating the accuracy of ReDi-Cell's cluttered reaction-diffusion computation.$k = 1.4^{12}s^{-1}$, $\Delta t = 10^{-14}s$, $D = 1.0\mu m^2/s$, $\Delta x = 10^{-6}m$

In all three sets of experiments run so far, we find a good agreement between ReDi-Cell's results and VCell's results. The concentration trajectories for VCell and ReDi-Cell are nearly identical in all three experiments. Very little variation is shown between the two algorithms. Heat map results also show good agreement. The shapes of the ReDi-Cell and VCell concentration profiles in all three experiments show good agreements between each other. The values displayed in the heat map also show a very good agreement. In summary, it can be seen that ReDi-Cell is an accurate simulation of reaction-diffusion systems.

4.5 Summary

In theory VCell could have been used to perform the experiments in this thesis, albeit at a reduced speed. When benchmarked using the validation experiments it was found that ReDi-Cell performed approximately 24x faster than VCell. It may be possible to port VCell to the CUDA architecture used by ReDi-Cell, but as VCell is closed source it is impossible for an end user to do so. In contrast, ReDi-Cell is open source allowing users to change the source code to take advantage of any future developments in computing architecture. In addition, the client/server nature of VCell means that jobs are scheduled before execution, possibly incurring an additional delay. Finally, ReDi-Cell uses the dedicated VisIt toolkit which enables results to be presented in a variety of ways not possible with the VCell viewer, such as 3D "volume-slices" of the simulation space.

In this chapter, we introduced ReDi-Cell, the simulation package that will be used to generate the results in chapters 5, 6 and 7 of this thesis. ReDi-Cell was developed to allow rapid simulation of cell dynamics that scales to the user's hardware. Abstractions for cellular pathways that will be used in simulations in this thesis, along with justifications for their choice were also presented. ReDi-Cell's GPGPU architecture, choice of numerical methods and visualisation capabilities, provided by VisIt, were also detailed. Finally, ReDi-Cell's

accuracy was verified by comparing concentration trajectories and heatmaps of concentration against VCell, a known-good simulation. ReDi-Cell will be used in Chapters 5, 6 and 7 to generate the results for this thesis.

CHAPTER 5

Quantifying the response of simulated biological pathways to common bacterial shapes

In this chapter, we detail three sets of experiments that we run with Reaction-Diffusion Cell (ReDi-Cell). The first is focused on validating ReDi-Cells implementation against the state of the art, Virtual Cell Simulation Package (VCell). The validation experiments use two different prototypes: diffusion only and $A + B \rightarrow C$. A summary of the reaction systems is shown in Table 4.1. In the second set, we investigate the behaviour of abstracted cellular pathways in different cell shapes. In the third set we investigate the impact of clutter on the same pathways. For a discussion of the choice of diffusion coefficients see appendix 1. Experiments take place in a variety of cluttered and uncluttered environments. The results take the form of concentration heat maps and concentration against time graphs.

Graphs of concentration against time are used when studying the dynamics of a reaction in a single volume. Heatmaps are used when measuring the concentration in many volumes at once, at a single point in time. We exploit this property to measure the effect of space on the distribution of concentration. Many heatmaps recorded at sequential times can be stitched together into an animation to show the same information as a concentration graph, but for multiple volumes at once. Both concentration against time graphs and heatmaps are common in cell simulation, with VCell.(against which we validate our results) using these measurements.

Table 5.1: Geometry of the cell shapes used in these experiments

Cell	Geometry
Coccus	Diameter: $50\mu m$
Bacillus	Diameter: $25\mu m$ Length: $150\mu m$
Spirillum	Diameter: $28\mu m$ Length: $168\mu m$ Winding diameter: $16\mu m$

Table 5.2: Description of the BMR system. The parameters used in the experiments are shown in the captions of the figures depicting experimental results. Initial positioning described in results. Concentration is measured in arbitrary units. Neumann boundary conditions. Parameters common to all BMR reactions: $\Delta t = 10^{-14}s$, $\Delta x = 10^{-6}m$, $k = 1.4^{12}s^{-1}$

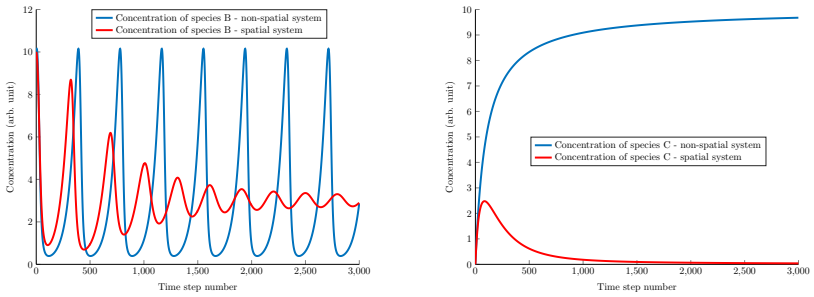
System	Reaction Equation	Initial Conditions
Diffusion		A:10
$A + B \rightarrow C$	$\frac{dA}{dt} = -k * A * B$ $\frac{dB}{dt} = -k * A * B$ $\frac{dC}{dt} = k * A * B$	A:10 B:10 C:0

Table 5.3: Description of the LVD system. The parameters used in the experiments vary and so are shown in the captions of the figures depicting experimental results. Neumann boundary conditions. Parameters common to all LVD reactions: $\Delta t = 10^{-14}s$, $\Delta x = 10^{-6}m$, $\alpha = 1.5 * 10^{12}$, $\beta = 1 * 10^{12}$, $\gamma = 3 * 10^{12}$

System	Reaction Equation	Initial Conditions
LVD	$\frac{dA}{dt} = A(\alpha - \beta * B)$ $\frac{dB}{dt} = -B(\gamma - \delta * A)$	A:10 B:10

5.1 The importance of spatial simulation

Having discussed the importance of spatial simulation, we show the impact that removing the spatial component has on a reaction-diffusion system. The system used in this experiment is the LVD system detailed in Table 5.3. To demonstrate this, we run two experiments. In the first experiment, we turn off diffusion allowing only reaction, thus making the experiment non spatial. In the second experiment we allow for both reaction and diffusion in a *Coccus*. The results of this experiment are shown in Figure 5.1.



(a) Concentration of *Prey* species in the sub-sub-volume that the chemical was initially confined to, using the LVD system. $D = 1.0\mu m^2/s$

(b) Concentration of species C in the sub-volume that the chemical was initially confined to, using the BMR system. $D = 0.02\mu m^2/s$

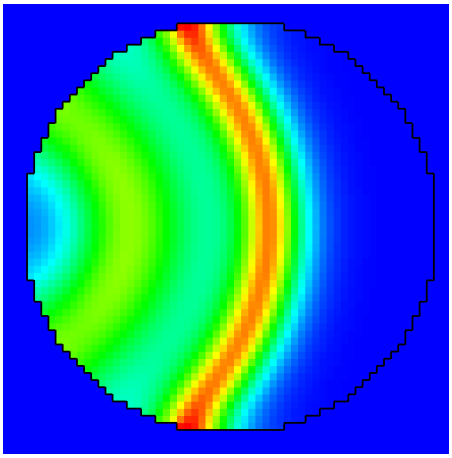
Figure 5.1: A comparison of spatial and non spatial systems in ReDi-Cell using both the LVD and BMR reaction systems. Concentration is measured in the sub-volume that the chemical was initially confined to. With both reaction systems it can be seen that there is a large difference in results between spatial and non spatial systems. Both the final value of the concentration and the trajectory of the reaction over time vary substantially.

5.2 LVD wavefronts in natural cell shapes

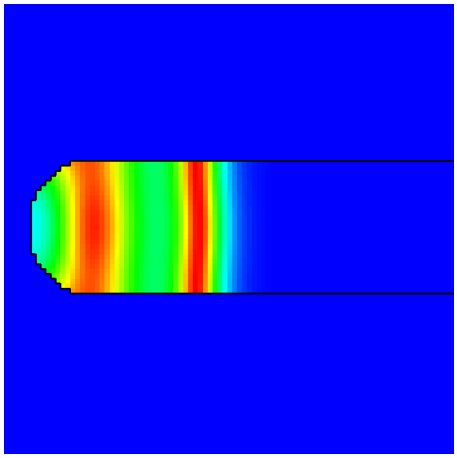
This set of experiments measures the effect of cell shape on wavefront properties, specifically wavefront shape and wavefront concentration. The system used in this set of experiments is the LVD system detailed in Table 5.3.

The results of this set of experiments are shown in Figure 5.2, Figure 5.3 and Table 5.4. The ReDi-Cell reaction vessels are approximations of real cell shapes. Three experiments were run, one in each cell shape. The volumes of

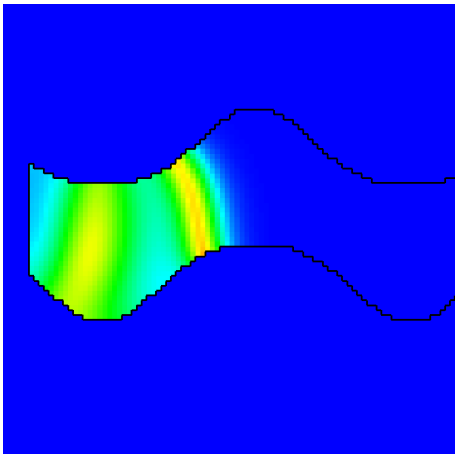
each cell type and initial concentrations of chemicals are kept identical. The initial positions of chemical species are also kept analogous; directly next to the membrane at the far left hand side of each cell.



(a) Concentration of Species A in a *Coccus*

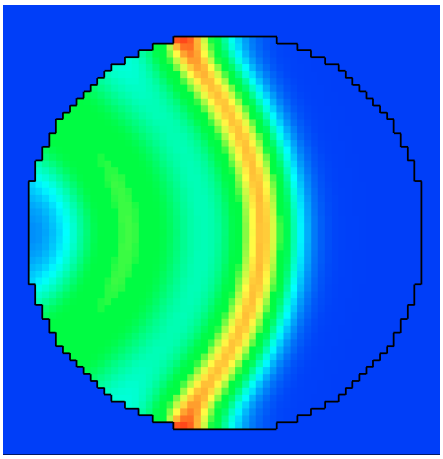


(b) Concentration of Species A in a *Bacillus*

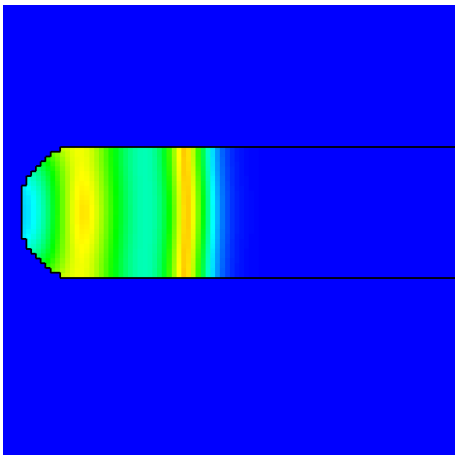


(c) Concentration of Species A in a *Spirillum*

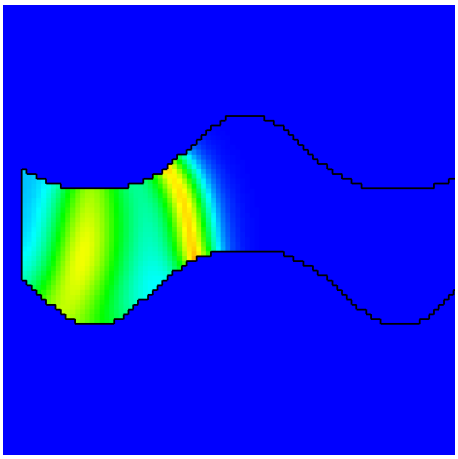
Figure 5.2: LVDwavefronts in natural cell shapes after 1000 time steps, simulated using ReDi-Cell and rendered in VisIt. We see the effect of running an LVD system in model natural cell shapes. In this figure each concentration gradient has its own scale. The shapes of the wavefronts are different in all three examples. The *Coccus'* wavefront is circular, whilst the *bacillus'* wavefront is almost flat. The *sprillum'*s wavefront is unlike the others in that the shape of the wave front is asymmetrical. $D = 1.0\mu m^2/s$



(a) Concentration of Species A in a *Coccus*



(b) Concentration of Species A in a *Bacillus*



(c) Concentration of Species A in a *Spirillum*

Figure 5.3: LVD wavefronts in natural cell shapes after 1000 time steps. These results differ from those in Figure 5.3; in Figure 5.2 the heatmap scales are different in all three cells, but in this result the heatmap scale is the same in all three cells so that concentrations can be compared across cells. Now that the heatmaps have been normalised it can be seen that in this central slice cell shapes induce differing concentrations. If we were to allow the wavefront of the *coccus* to continue past 1000 timesteps we would see that it reaches the other end of the cell and then collapses. $D = 1.0\mu m^2/s$

Next, we examine the effect that cell shape has on average wavefront concentration. The results are recorded in Table 5.4.

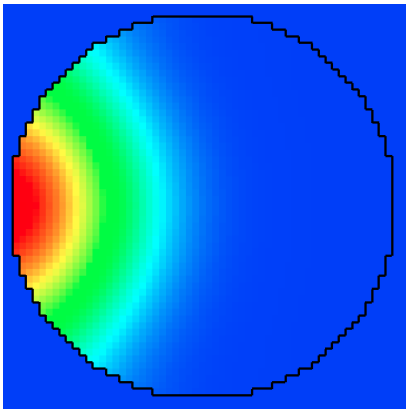
Table 5.4: LVD wavefronts in natural cell shapes. Concentration at the wavefront after 1000 time steps. The average concentration at the wavefront in all three bacterial shapes is different, agreeing with the results in Figure 5.3 $D = 1.0\mu m^2/s$

Shape	Concentration At The Wavefront
Coccus	1.06
Bacillus	1.38
Spirillum	2.61

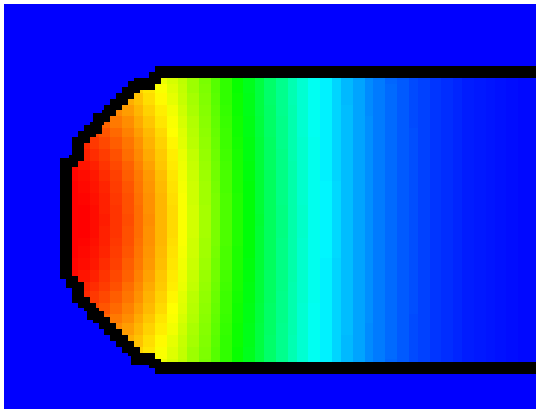
Different surface area to volume ratios may be responsible for the unique patterns of reflection inside virtual cells and so cause the difference in concentration gradients and wavefront shape between cell shapes seen in this section. In geometries with lines of symmetry, such as the *Coccus* and *Bacillus*, concentration waves are reflected in a uniform, symmetric fashion. The *Spirillum* is nonuniform in shape and thus reflections are not necessarily symmetric about the longest axis of the cell. This may produce the asymmetric wavefronts seen in the results. Changes in reflective properties may alter micro dynamics and so change the global response of the system. In addition, in cells of equivalent volume, signals in the simulated *Coccus* travel a greater fraction of the cell's length compared to the other cell types using the same timestep size. As the number of timesteps in each simulation is the same, this is equivalent to saying that wave in the *coccus* moves faster than in other cells.

5.3 BMR wavefronts in natural cell shapes

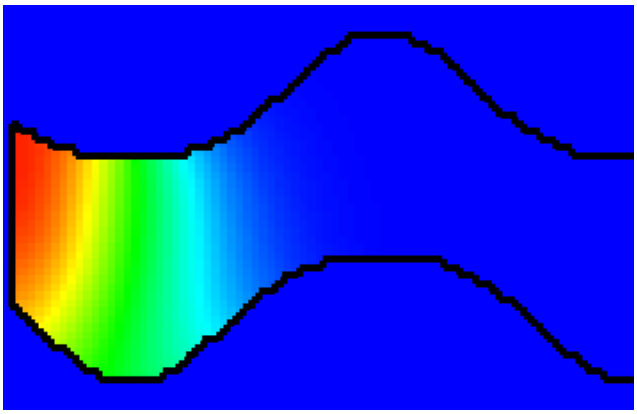
This set of experiments measures the effect of cell shape on wavefront properties within the BMR system (Table 5.2). The results of this set of experiments are shown in Figure 5.4 and Figure 5.5. Experimental set up is identical to the LVD experiments aside from the change in reaction system.



(a) Concentration of Species C in a *Coccus*

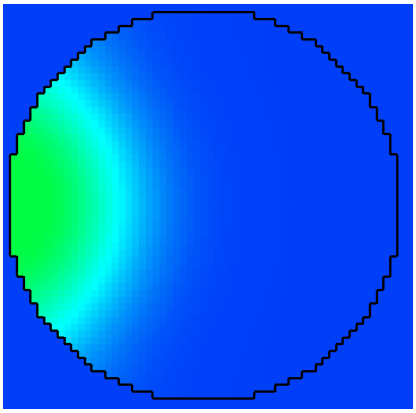


(b) Concentration of Species C in a *Bacillus*

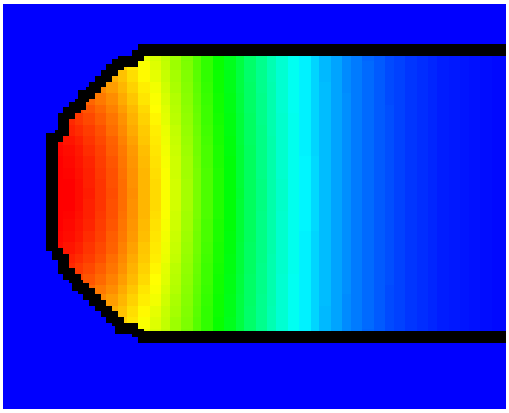


(c) Concentration of Species C in a *Spirillum*

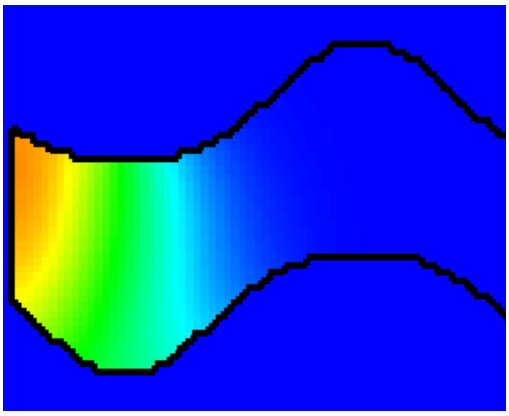
Figure 5.4: BMR wavefronts in natural cell shapes after 1000 time steps, simulated using ReDi-Cell and rendered in VisIt. We see the effect of running a BMRsystem in model natural cell shapes. Each concentration gradient has its own scale. The shapes of the wavefronts are different in all three virtual cells. The results for the BMR system mirror the LVD system. The *Coccus'* wavefront is circular, whilst the *Bacillus'* wavefront is almost flat. $D = 1.0\mu m^2/s$



(a) Concentration of Species C in a *Coccus*



(b) Concentration of Species C in a *Bacillus*



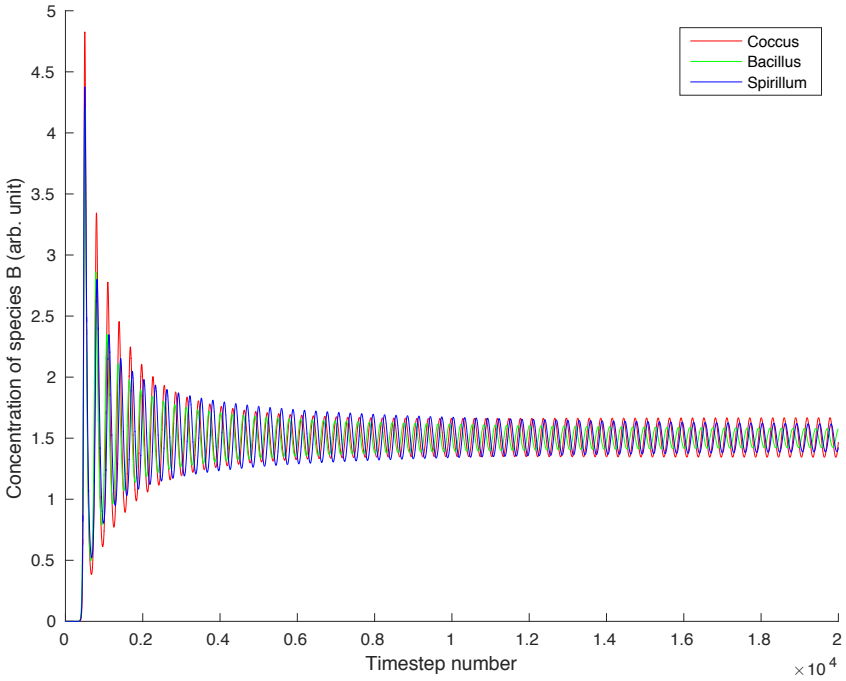
(c) Concentration of Species C in a *Spirillum*

Figure 5.5: BMR wavefronts in natural cell shapes after 1000 time steps, simulated using ReDi-Cell and rendered in VisIt. We see the effect of running a BMR system in model natural cell shapes. Heatmaps are normalised so that concentrations can be compared. $D = 1.0\mu m^2/s$

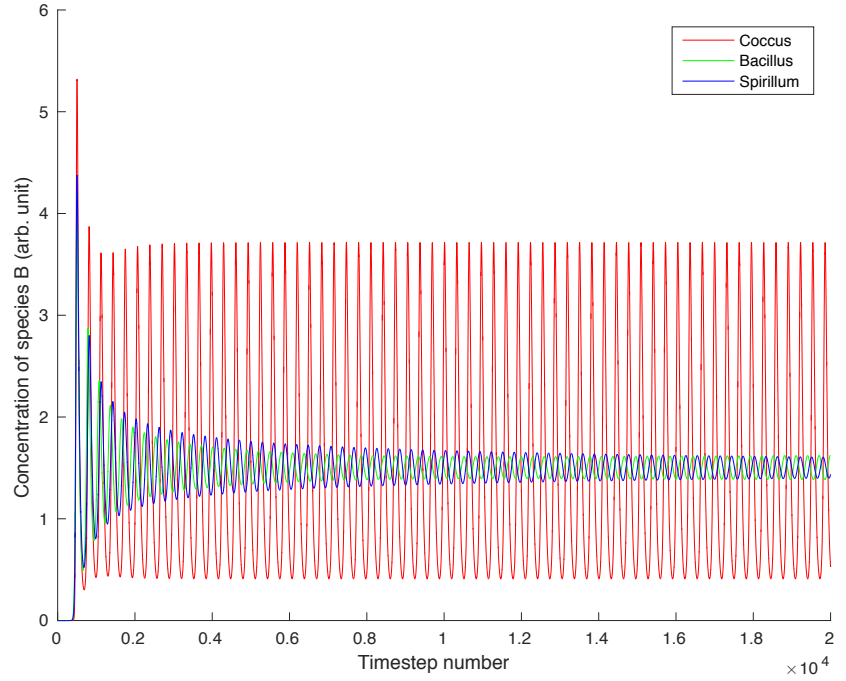
As with the LVD system, we see that the BMR experiments systems with symmetrical geometries (the *Coccus* and *Bacillus*) produced symmetrical waves of concentration. As with the LVD system, the *Spirillum* produces an asymmetrical leading edge.

5.4 The effect of initial positioning on pathway dynamics

This set of experiments examines the effect that initial positioning has on virtual cell reaction dynamics. We first determine if initial position has *any* effect on reaction dynamics. We compare two scenarios. In the first scenario, the component species of the pathway are positioned at the far left hand side of the cell and time series of their concentrations measured at the far right hand side of the cell. In the second scenario, the component species are positioned in the centre of the cell, and time series of their concentrations are once again measured at the far right of the cell. Each initial position scenario is repeated for all three cell shapes and both reaction systems.

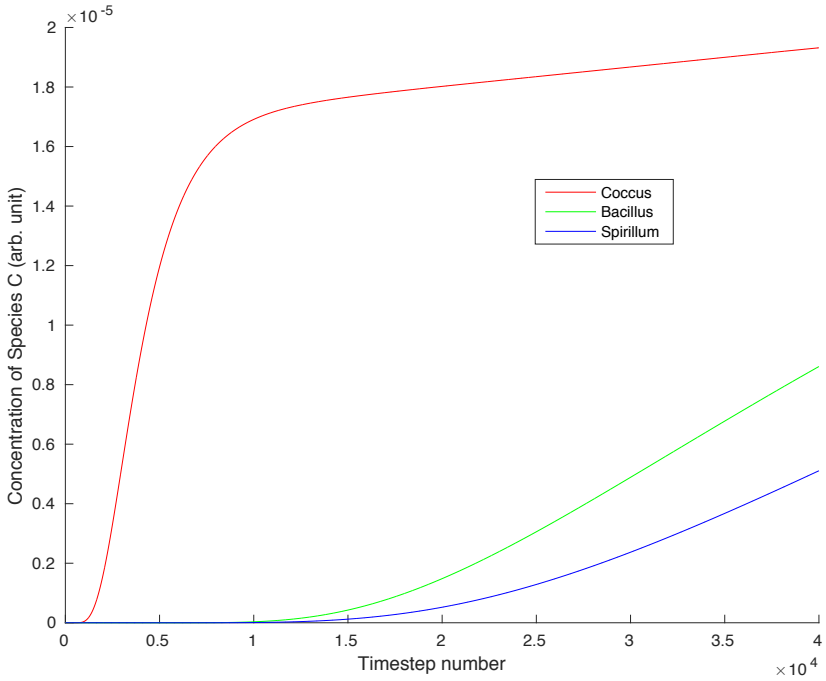


(a) Concentration of Species B when pathway species are initially positioned at the extreme left of the cell

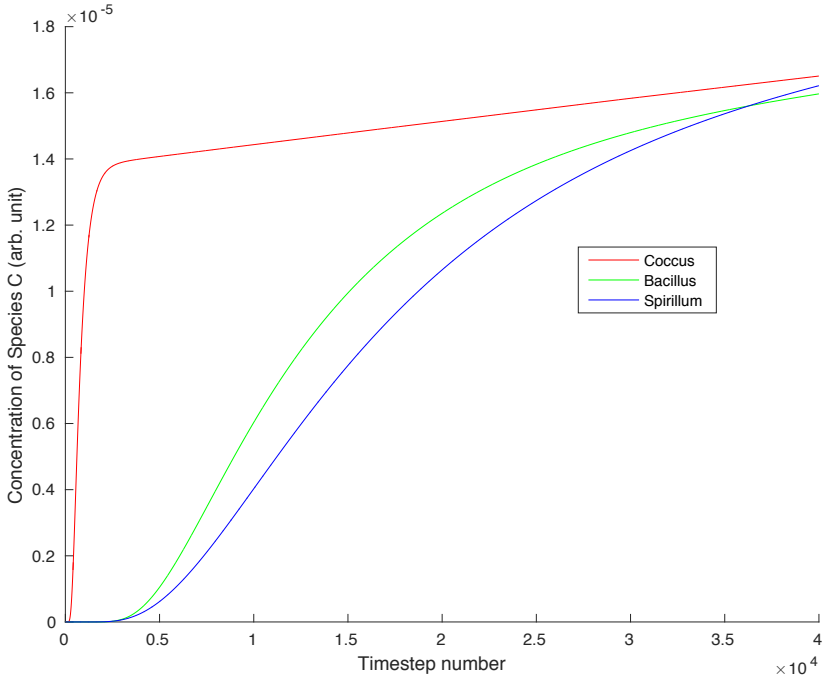


(b) Concentration of Species B when pathway species are initially positioned at centre of the cell

Figure 5.6: Time series of species B concentration in both initial position configurations. Changing the initial position of the pathway constituents alters cellular dynamics. $D = 1.0\mu m^2/s$



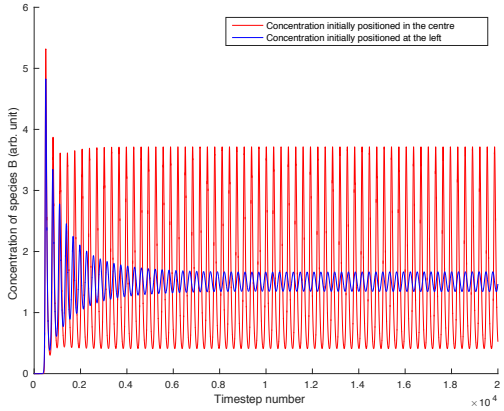
(a) Concentration of Species C when pathway species are initially positioned at the extreme left of the cell



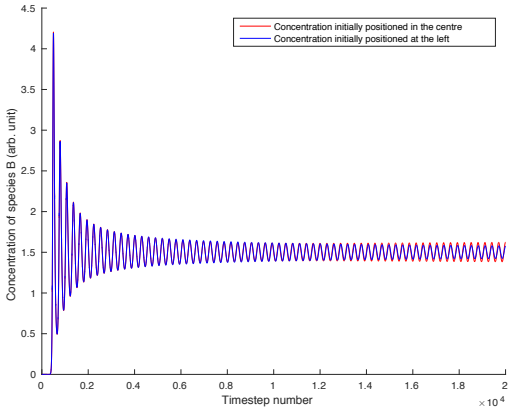
(b) Concentration of Species C when pathway species are initially positioned at centre of the cell

Figure 5.7: Time series of species C concentration in both initial position configurations. Changing the initial position of the pathway constituents alters cellular dynamics. $D = 1.0\mu m^2/s$

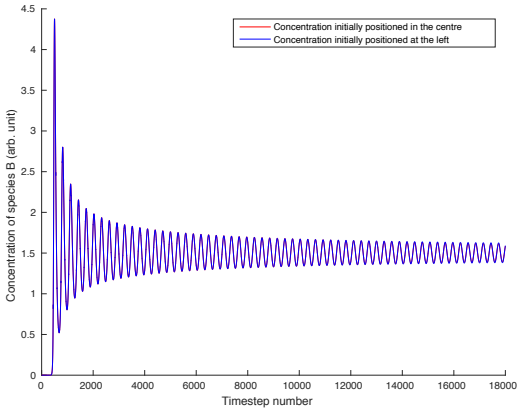
Having seen that virtual pathway dynamics can vary between initial configurations we now examine differences in concentration time series within the same cell shapes. These results can be seen in Figure 5.8 and Figure 5.9.



(a) Concentration of Species B in a virtual coccus from both initial configurations

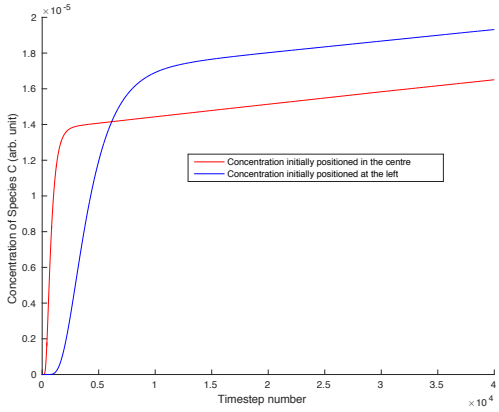


(b) Concentration of Species B in a virtual bacillus from both initial configurations

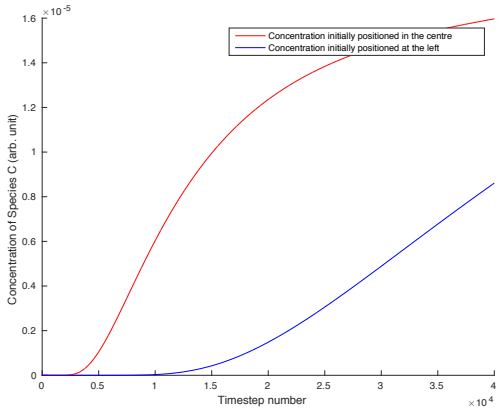


(c) Concentration of Species B in a virtual spirillum from both initial configurations

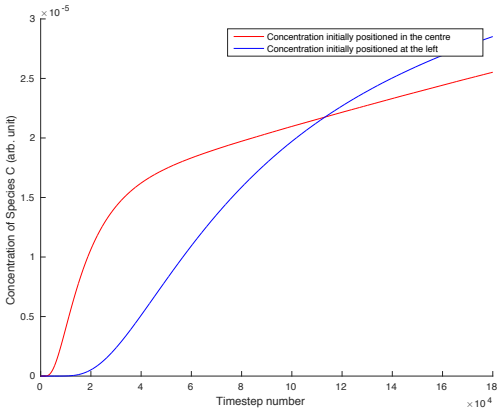
Figure 5.8: Time series of species B concentration in both initial position configurations. $D = 1.0\mu m^2/s$



(a) Concentration of Species C in a virtual coccus from both initial configurations



(b) Concentration of Species C in a virtual bacillus from both initial configurations



(c) Concentration of Species C in a virtual spirillum from both initial configurations

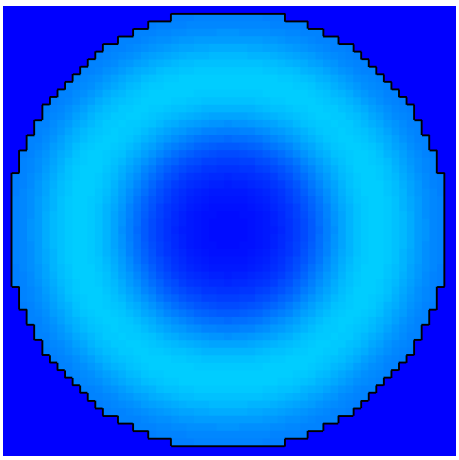
Figure 5.9: Time series of species C concentration in both initial position configurations. Changing the initial position of the pathway constituents alters cellular dynamics. $D = 1.0\mu m^2/s$

Results for both reaction systems need to be presented in different ways. In simulations run in the BMR system, differences in the time series between initial configurations can easily be seen. In the LVD system the differences between central and left initial configurations can only be easily seen in the *Coccus*. In order to show the difference between between the *Bacillus* and *Spirillum* a second order exponential curve of the form $ae^{bx} + ce^{dx}$ is fit to the peaks of the LVD signals. Peaks before the maximum global concentration are removed. The parameters of these four curves can be seen in Table 5.5

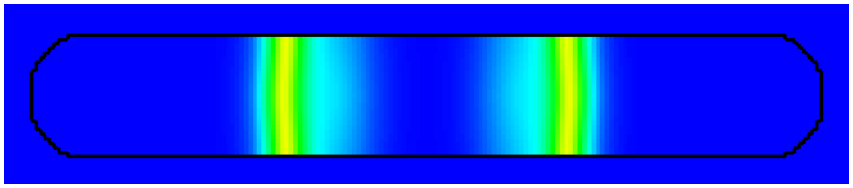
	Bacillus left initial configuration	Bacillus, central initial configuration	Spirillum, left initial configuration	Spirillum, central initial configuration
a	0.43	0.54	0.51	0.49
b	-0.0004	-0.0004	-0.0004	-0.0004
c	1.64	1.58	1.71	1.72
d	$-1.94 * 10^{-6}$	$1.35 * 10^{-6}$	$-2.9 * 10^{-6}$	$3.765 * 10^{-6}$

Table 5.5: Constants resulting from fitting a second order exponential of the form $ae^{bx}+ce^{dx}$ to peaks in the time series of LVD simulations. LVD simulations are of the *Bacillus* and *Spirillum* virtual cells in both left and centre initial configurations. $D = 1.0\mu m^2/s$

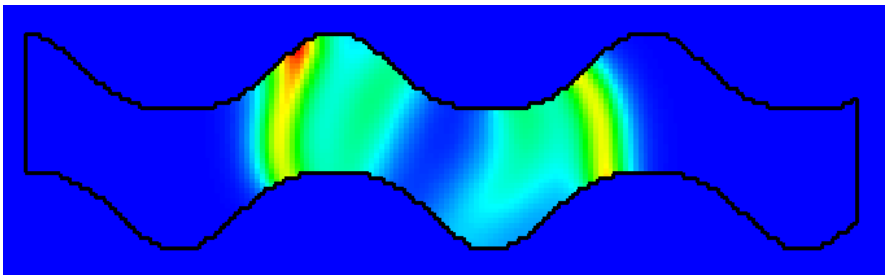
Finally we show that the moving the initial position of pathway constituents alters the nature of the concentration gradients observed in the virtual cells Figure 5.10 and Figure 5.11.



(a) Concentration of Species B in a virtual coccus from both initial configurations

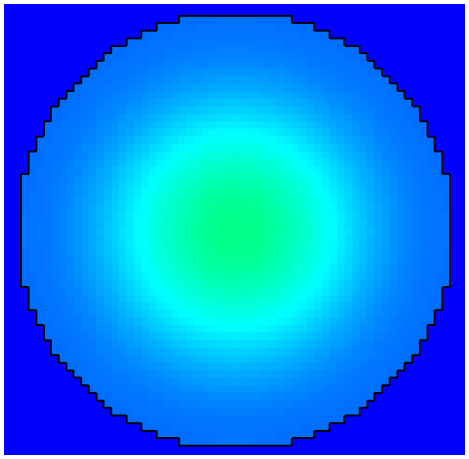


(b) Concentration of Species B in a virtual bacillus from both initial configurations

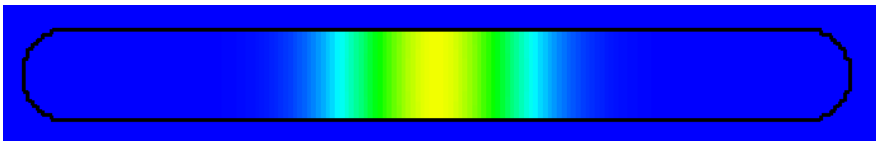


(c) Concentration of Species B in a virtual bacillus from both initial configurations

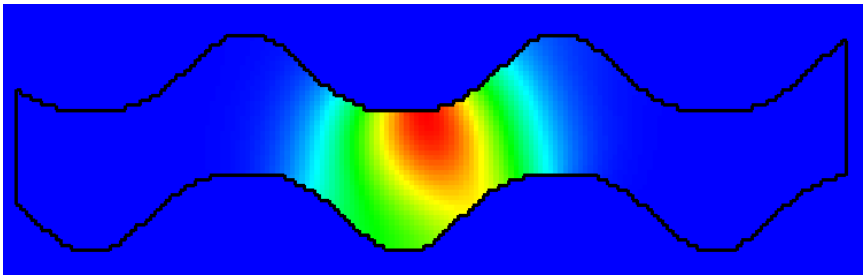
Figure 5.10: LVDwavefronts in natural cell shapes after 1000 time steps, simulated using ReDi-Cell and rendered in VisIt. We see the effect of running an LVD system in model natural cell shapes when the concentration is initially positioned at the centre. The concentration gradients show a marked difference to those in Figure 5.3 where concentration is initially located at the left of the cell. All heat maps in this figure are normalised. $D = 1.0\mu m^2/s$



(a) Concentration of Species C in a *Coccus*



(b) Concentration of Species C in a *Bacillus*



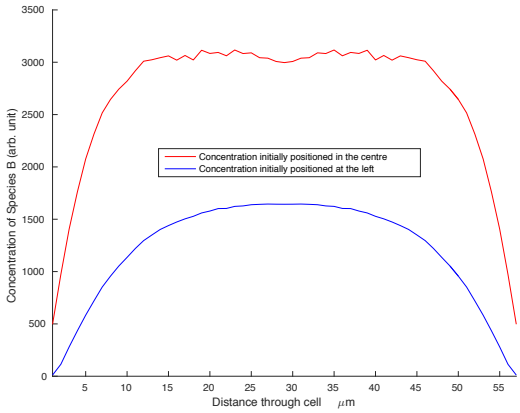
(c) Concentration of Species C in a *Spirillum*

Figure 5.11: BMRwavefronts in natural cell shapes after 1000 time steps, simulated using ReDi-Cell and rendered in VisIt. We see the effect of running a BMR system in model natural cell shapes when the concentration is initially positioned at the centre. The concentration gradients show a marked difference to those in Figure 5.5 where concentration is initially located at the left of the cell. All heat maps in this figure are normalised. $D = 1.0\mu m^2/s$

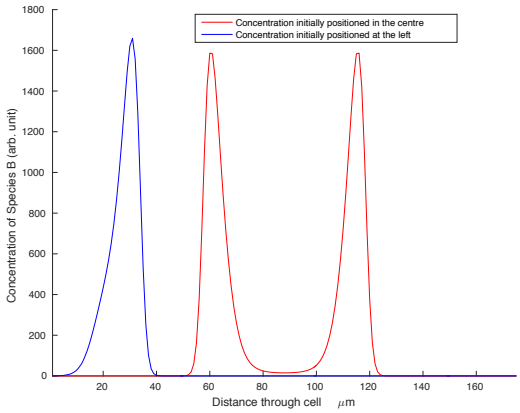
5.5 LVD species distribution

Next we examine the distribution of pathway chemical species in the three virtual cell shapes. Species distribution is measured as follows. The 3D cell is divided into 2D slices along the cell's longest axis, with each slice having a thickness equal to one simulation unit length, Δx. The total amount of a given species in each slice is found by summing over each sub-volume in the slice. Finally, the total amount of a particular species in each slice is plotted against the position of that slice along the cell.

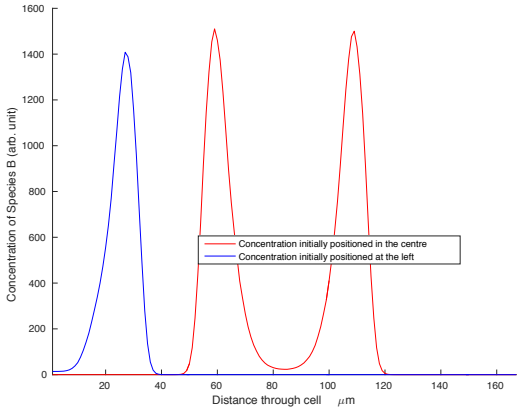
The LVD system detailed in Table 5.3 is run in the three different cell shapes, with both starting positions. The species distribution profile is recorded after 1000 timesteps. The species distributions are shown in Figure 5.12.



(a) Distribution of Species A in a *Coccus*



(b) Distribution of Species A in a *Bacillus*

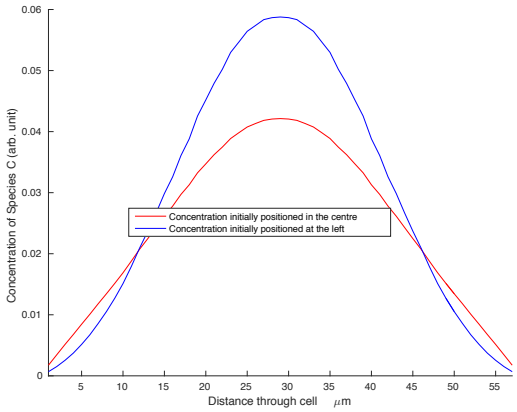


(c) Distribution of Species A in a *Spirillum*

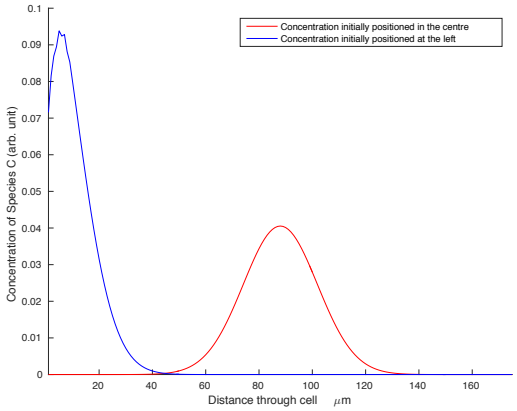
Figure 5.12: LVD concentration distribution after 1000 timesteps. A clear difference in concentration profile can be seen between the three virtual cells and the different initial positions. We see several differences between the graphs in Figure 5.12. Different cellular geometries result in different concentration profiles and peak concentrations. The shapes of the concentration profiles are different in each example $D = 1.0\mu m^2/s$

5.6 BMR species distribution

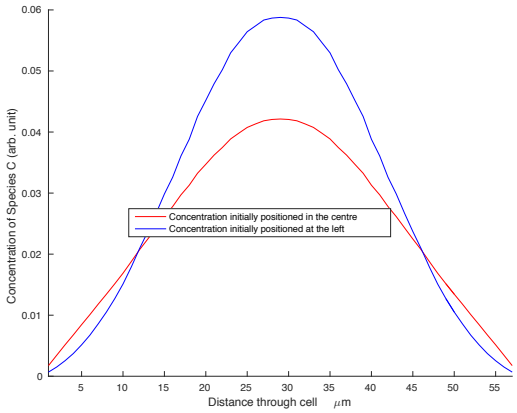
This set of experiments repeats those in the previous section, but substitutes the LVD system for the BMR system. The BMR system detailed in Table 5.2 is run in the three different cell shapes. The species distribution profile is recorded after 1000 timesteps. The species distributions are shown in Figure 5.13.



(a) Distribution of Species C in a *Coccus*



(b) Distribution of Species C in a *Bacillus*



(c) Distribution of Species C in a *Spirillum*

Figure 5.13: BMR concentration distribution after 1000 timesteps. The different cell shapes impose differing peak concentration and concentration distributions on the simulated BMR pathway. Concentration profiles differ between cells and within the same cell when the starting location is changed. $D = 0.5\mu m^2/s$

5.7 The sensitivity of the Turing system to intracellular structures

In this final section we investigate the spatial sensitivity of the Turing system and compare it to the BMR and LVD systems. This system, along with a Java simulator to demonstrate has been published previously [Kondo and Miura, 2010].

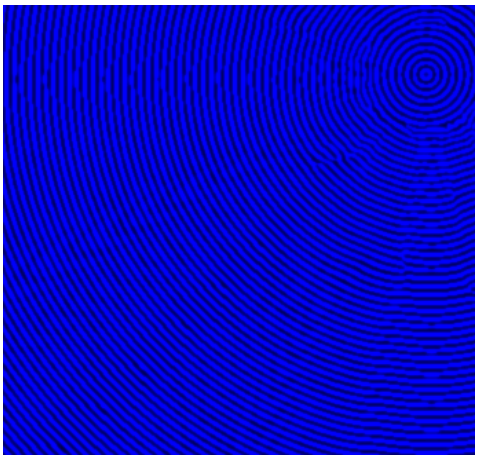
$$\frac{\partial u}{\partial t} = F(u,v) - d_u U + D_u \Delta u \tag{5.1}$$

$$\frac{\partial v}{\partial t} = G(u,v) - d_v V + D_V \Delta v \tag{5.2}$$

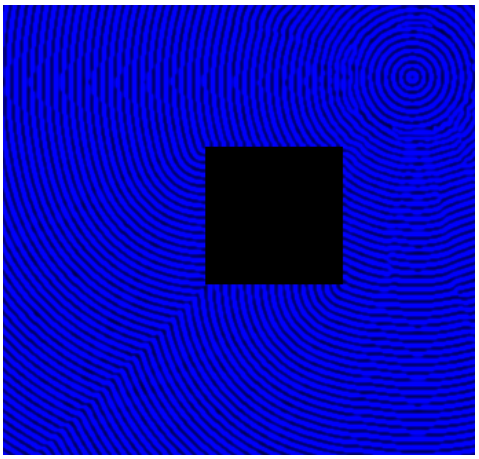
$$0 \leq F(u,v) = a_u u + b_u v + c_u \leq Fmax \tag{5.3}$$

$$0 \leq G(u,v) = a_v u + b_v v + c_v \leq Gmax \tag{5.4}$$

Here a single structure within the virtual cell causes discontinuities in the generated pattern.



(a) Concentration of Species V in a cluttered environment



(b) Concentration of Species V in an un-occupied environment

Figure 5.14: Concentration map of species V in a Turing system after 10000 time steps. The concentration of species V is proportional to the intensity of the blue. $a_u = 0.08$, $b_u = -0.08$, $c_u = 0.04$, $d_u = 0.03$, $reaction\ rate_u = 10^{12}$, $D_u = 0.02$, $a_v = 0.1$, $b_v = 0.0$, $c_v = -0.15$, $d_v = 0.08$, $reaction\ rate_v = 10^{12}$, $D_v = 0.5$

5.8 Analysis of simulations

We have performed simulations that show how the behaviour of model cellular pathways changes between three common bacterial cell shapes, and how incorporating space into the simulation of a reaction system can alter the way in which signals propagate through a virtual cell. In doing so, we show that model

cellular pathways are sensitive to the shape and the internal geometry of the cell. Most interestingly, we show that the two pathways have different responses to the same changes in morphology. In the LVD system, concentration gradients within the virtual cells are quite different to one another, but the level of concentration behind the wavefront is similar in all three cells. The BMR system on the other hand has nearly identical patterns of concentration gradient across all three virtual cells, but the relative concentrations are different. Whilst smooth predictable distributions of concentration are found within the BMR system, the LVD system displays a far more jagged appearance. As a final note on pathway specific sensitivity, we see that in both of the oscillatory virtual pathways (the Turing system and the LVD system) absolute concentrations remain relatively unaffected, although we do see that structures within the cell have the potential to alter wavefront dynamics.

Finally, the wavefront shapes, for the same virtual cell, in the different reaction systems are quite similar. This would indicate the independence of wavefront shape with respect to reaction system.

CHAPTER 6

Quantifying the responses of simulated biological pathways to internal cellular organisation

The simulations that we perform in this chapter investigate the effect that three spatial variables (Nuclear to Cytoplasm (NC) ratio, nuclear position and substructure distribution - discussed in chapter 2) have on the behaviour of model cellular pathways.

All of the simulations in this work take place inside a model *Coccus* of constant volume. Inside this virtual cell, we place impermeable spherical structures that represent cell nuclei (in the case of measuring the effect of NC ratio and nuclear position) or organelles (in the case of measuring the effect of substructure distribution). For each spatial variable, both the Lotka-Volterra-Diffusion (LVD) and Bi-Molecular Reaction (BMR) pathways are simulated. Table 6.3 shows a summary of the model pathways and their parameters used in this experiment.

For each spatial variable, we experiment with a range of Structure to Cytoplasm (SC) ratios. These range from 0% to 80% when investigating NC ratios and nuclear position, and from 0% to 50% when investigating the distribution of substructure. Past 50 % occupancy it becomes challenging to place non-overlapping structures within the virtual cytoplasm. When investigating nuclear size the position of the nucleus is kept constant. On the other hand when investigating nuclear positioning and substructure distribution, many simulations are performed - 100 simulation runs at each SC ratio. When experimenting with nuclear position, each simulation run sees the nucleus placed at a different location within the cell. When experimenting with structure distribution, each simulation run has a different random positioning of model organelles.

Table 6.1: Summary of Simulations Performed

Variable	SC ratio range	Measurements	Systems used
Nuclear size	0-80%	1	BMR, LVD
Nuclear position	0-80%	100	BMR, LVD
Substructure distribution	0-50%	100	BMR, LVD

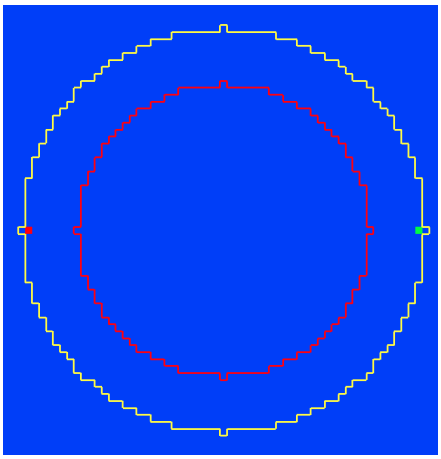


Figure 6.1: ReDi-Cell example cell environment showing the initial location of all species (red square) origin and measurement point (green square). An impermeable example nucleus is represented by the red shape and the cell boundary is shown in yellow. In all experiments, the origin and measurement locations are kept constant. Nucleus sizes and positions change.

A summary of the simulations performed can be seen in Table 6.1.

In all simulations, the constituents of the virtual pathways are placed in the same location on the left hand side of the model cell. We call this location the Original Point (OP). Exactly opposite this point, on the right hand side of the cell, we measure the system by recording time series of constituent concentrations. This location is known as the Measurement Point (MP). The locations of these two points within the cell can be seen in Figure 6.1.

As stated above, we record chemical concentration time series within the virtual cell at MP. These time series can be displayed independently of one another or plotted against each in the form of phase-space diagrams. Phase space diagrams show all possible states of the system independently of time. Phase space diagrams are often used to study the behaviour of oscillators, such as the LVD system.

Time series show concentration information at one point (MP) in the cell over a specific time interval. Conversely, heatmaps show concentration at many points in the cell but only at a single point in time. The heatmaps in this work show concentration information in a 2D slice through the virtual cell. This 2D slice is taken at the same plane as the source and measurement sites. Heatmaps are commonly used to represent concentration gradients within cell simulations. In this work we refer to Simulation sets or *SS* to identify results. Often one figure is used to compare the results of many different simulation setups. In other cases one simulation setup is analysed in a few different figures, for example one figure is used to record many timeseries and another to record the peak concentrations of those timeseries. The different simulation setups can be found in Table 6.2.

Table 6.2: List of all simulation setups used in this chapter

Simulation Set (SS)	Spatial Variable Investigated	Reaction System	NC/SC Ratios Simulated	Number of Simulations at Each NC/SC Ratio	Type of Measurement
SS1	NC Ratio	LVD	0,50,80	1	Timeseries
SS2	NC Ratio	LVD	0,10,20,30,40,50,60,70,80	1	Timeseries
SS3	NC Ratio	LVD	0,10,50	1	Heatmap
SS4	NC Ratio	BMR	0,50,80	1	Timeseries
SS5	NC Ratio	BMR	0,10,20,30,40,50,60,70,80	1	Timeseries
SS6	NC Ratio	BMR	0,10,50	1	Heatmap
SS7	Nuclear Position	LVD	10	4 (in selected positions)	Timeseries
SS8	Nuclear Position	LVD	50	4 (in selected positions)	Timeseries
SS9	Nuclear Position	LVD	10,20,30,40,50,60	100 (in random positions)	Timeseries
SS10	Nuclear Position	LVD	10	2 (in selected positions)	Heatmap
SS11	Nuclear Position	BMR	10	4 (in selected positions)	Timeseries
SS12	Nuclear Position	BMR	50	4 (in selected positions)	Heatmap
SS13	Nuclear Position	BMR	30,40,50,60,70,80	100 (in random positions)	Timeseries
SS14	Nuclear Position	BMR	10	2 (in selected positions)	Heatmap
SS15	Substructure distribution	LVD	20	3 (in selected configurations)	Timeseries
SS16	Substructure distribution	LVD	50	2 (1 selected, 1 random configuration)	Timeseries
SS17	Substructure distribution	LVD	10,20,30,40,50	100 (in random positions)	Timeseries
SS18	Substructure distribution	BMR	20	3 (in selected configurations)	Timeseries
SS19	Substructure distribution	BMR	50	2 (1 selected, 1 random configuration)	Timeseries
SS20	Substructure distribution	BMR	10,20,30,40,50	100 (in random positions)	Timeseries

Table 6.3: Description of the systems used in these simulations. Neumann boundary conditions. $\Delta t = 10^{-14}s$, $\Delta x = 10^{-6}m$, $\alpha = 1.5 * 10^{12}$, $\beta = 1 * 10^{12}$, $\gamma = 3 * 10^{12}$, $k = 1.4^{12}s^{-1}$

System	Reaction Equation	Initial Conditions	Diffusion constant of all chemicals ($\mu m^2/s$)
BMR	$\frac{dA}{dt} = -k * A * B$ $\frac{dB}{dt} = -k * A * B$ $\frac{dC}{dt} = k * A * B$	$A = 10$ $B = 10$ $C = 0$	2.0
LVD	$\frac{dpred}{dt} = pred * (\alpha - \beta * prey)$ $\frac{dprey}{dt} = -prey(\gamma - \delta * pred)$	$pred = 10$ $prey = 10$	0.05

Table 6.4: Cell geometry used in these experiments

Cell	Geometry
Coccus	Diameter: $50\mu m$

The timeseries results in the sections dealing with nuclear position and volume distribution require additional explanation. The first is the "central timeseries". This is the timeseries of concentration that results from placing a single subvolume, of size ratio NC, (modelling a nucleus) in the centre of the cell, and then running the reaction system. The second is the "average position timeseries". A single subvolume (modelling a nucleus), of size ratio NC, in a random position inside the cell. For each of these placings one simulation is run, producing one timeseries. After many simulations have been run (and many timeseries have been produced) the timeseries are averaged. The averaging processes works by aligning all series so that they have the same start point, and them computing the average timeseries. The arrival time of this new average time series is calculated by averaging the arrival times. The "average distribution timeseries" describes behaviour for distributed volumes. For a given SC ratio we randomly place 5% ratio subvolumes inside of the cell that, summed, make up the entire SC ratio. For each of these placings one simulation is run, producing one timeseries. After many simulations have been run (and many timeseries have been produced) the timeseries are averaged. The averaging processes works in the same way as the "average position timeseries". By computing these timeseries we can compare the "average" behaviour of centrally placed volume, non

centrally placed volume and divided volume, while controlling for volume.

6.1 NC ratio alters virtual pathway behaviour

In this first set of simulations, we model the impact that NC ratio has on the properties of two abstract signalling pathways by simulating cells of different NC ratios.

First, we simulate 0%, 50% and 80% NC ratios. Next, a range of NC ratios from 0% to 80% are simulated; peak concentrations and arrival times of chemical species from this range of NC ratios are analysed. Finally, concentration gradients in 0%, 10%, 50% NC ratios are explored. This list of simulations is used for both LVD, BMR reaction systems.

6.1.1 The LVD system

Initially, we compare signalling behaviour in three virtual cells displaying NC ratios of 0%, 50% and 80%. These figures are representative of some types of cells. In particular, maturing *blast erythrocytes* take on NC values between 50% and 80% [Turgeon, 2005]. The 0% NC ratio cell is provided as a control. In all simulations of this type, the nuclei are positioned at the centre of the cell as depicted in Figure 6.1. In Figure 6.2 we observe the effect that NC ratio has on the behaviour of the LVD system.

As we discussed in chapter 2, NC ratios rarely remain constant throughout the life of a cell. We repeat *SS1* with NC ratios of 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70% and 80%. The results of these simulations are seen in Figure 6.3.

Figure 6.3 shows the effect that a range of NC ratios have on the LVD system in the form of a phase space diagram. The relationship between NC ratio and chemical concentration at the measurement site is complex. Concentration behaviour is further explored by plotting the peak concentration of *Predator* species against NC ratio in Figure 6.4.

The time of arrival of *Predator* species as a function of NC ratio is explored

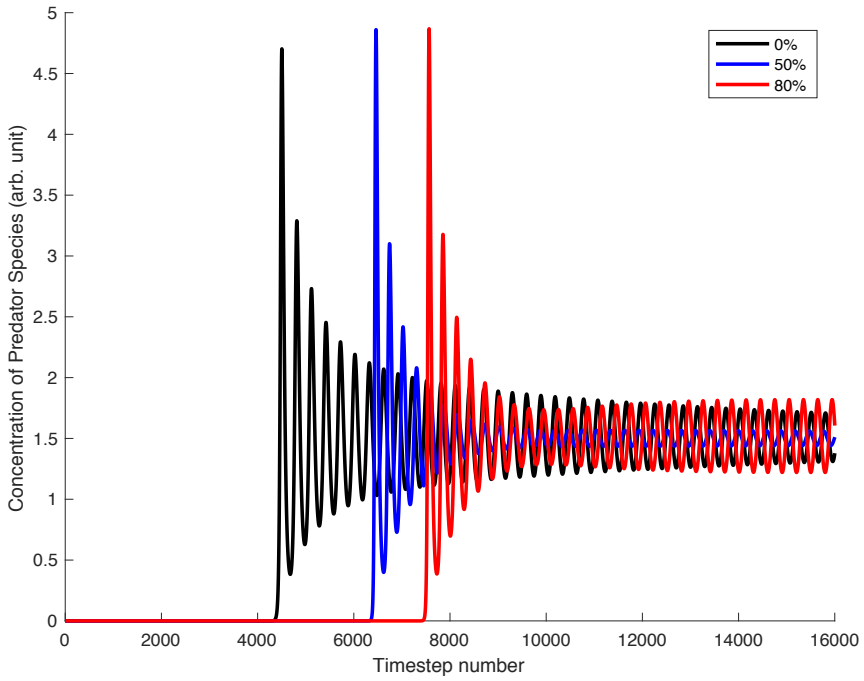


Figure 6.2: *SS1 results*:Time series of *Predator* species concentration. In the virtual cells with larger NC ratios the time of arrival of the *Predator* species is delayed. NC ratio alters the amplitude of the timeseries. It appears that different NC ratios cause different degree of damping, but without further analysis it is difficult to tell.

in Figure 6.5.

Having established that NC ratio affects arrival time it is evident that at the same timestep two cells with different NC ratios will display different concentration gradients, simply because the travelling wave of concentration will be in a different position. In SS3 we examine concentration gradients when the travelling wave of concentration has reached the same position in the virtual cell. This allows us to represent chemical processes *at the same stage* in cells of different NC ratios. With this approach it is possible to ask question such as: "What are the differences in concentration gradients (in cells of different NC ratio) when a signalling processes has just reached the target position?". Figure 6.6 shows concentration gradients in virtual cells of differing NC ratios when the wavefront has reached the same position in the cell.

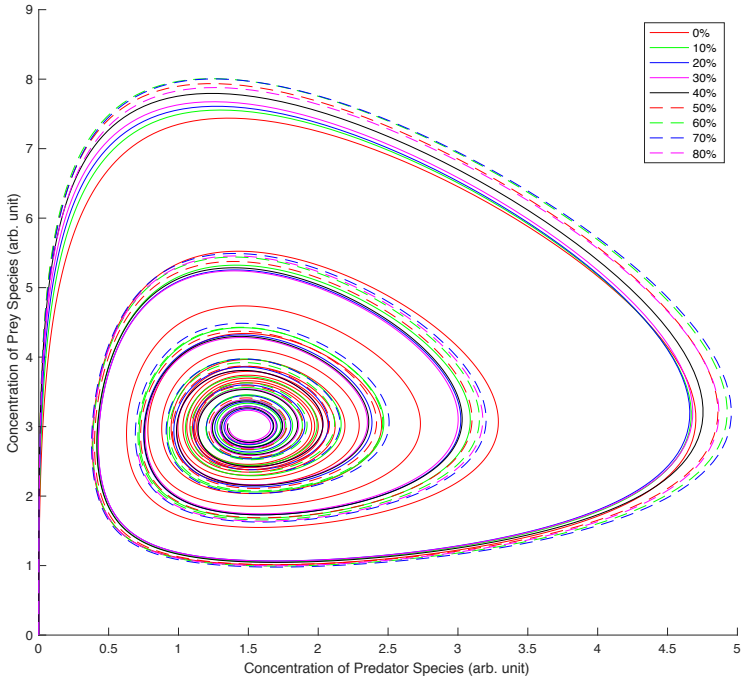


Figure 6.3: *SS2 results*:Phase diagram of species concentrations for a range of NC ratios. Concentration is measured at MP. A point in the phase space diagram is the value of predator and prey concentration at a particular point in time in the cell. A trajectory represents predator and prey concentration combinations through the simulation

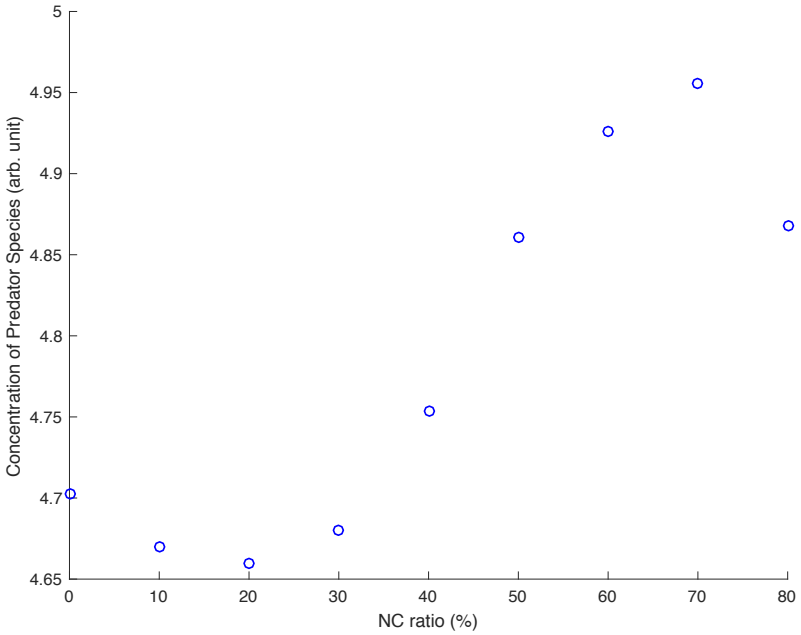


Figure 6.4: *SS 2 results*:Peak concentration of *Predator* species in virtual cells of different NC ratios. Peak concentration refers to the global maximum concentration over the entire timeseries of one experiment. This result shows the complex non-linear response of peak *Predator* concentration to NC ratio.

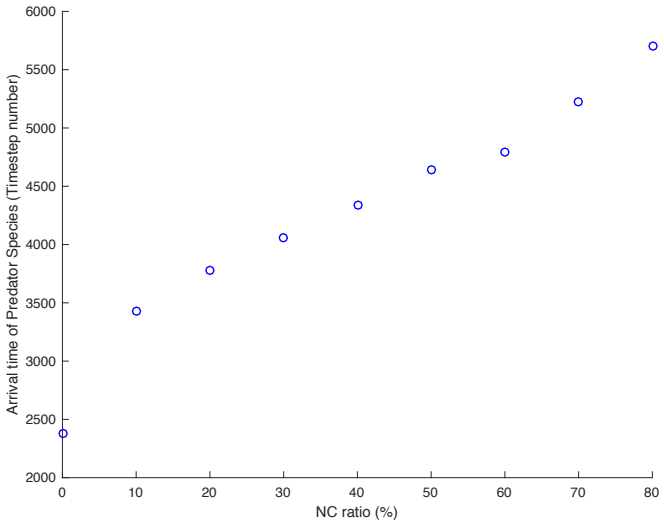
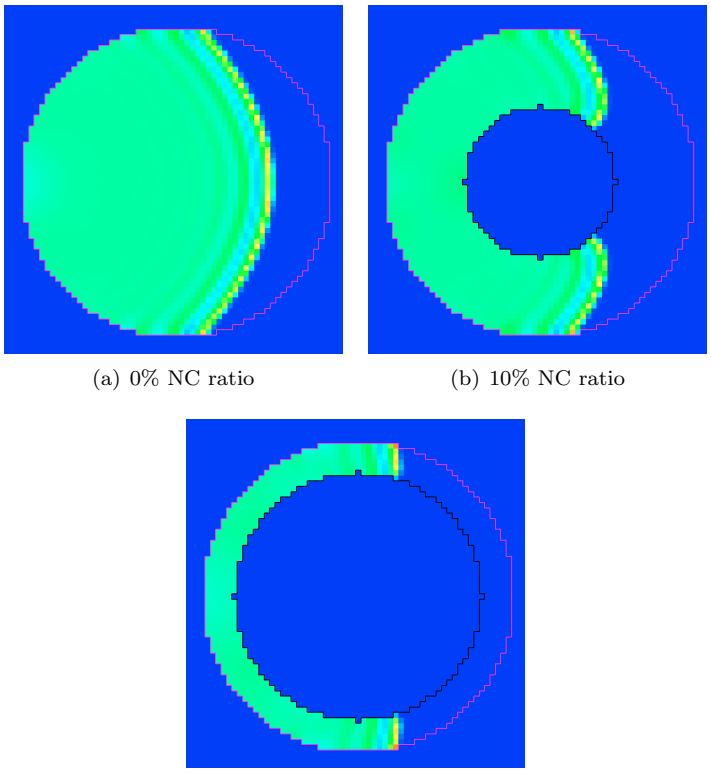


Figure 6.5: *SS 2 results*:Arrival time of *Predator* species in cells at different NC ratios. Arrival time is measured in the destination sub-volume. Arrival time is defined as the timestep number when the concentration at the measurement point becomes non zero. It can be seen that arrival time increases linearly with NC ratio.



(a) 0% NC ratio

(b) 10% NC ratio

(c) 50% NC ratio

Figure 6.6: *SS3 results*:Concentration gradients of *Predator* species in the LVD system at different NC ratios. Gradients captured when wavefronts reach the same position. Differing NC ratios induce different concentration gradients in the LVD system. Wavefront shape and peak concentration are different in all three examples

6.1.2 The BMR system

Having explored the response of the LVD system to NC ratio, we now explore the spatial properties of the BMR system. We repeat the simulations from SS1, except substituting the LVD system for the BMR system. The time series from the BMR reactions are shown in Figure 6.7.

Next, repeating SS2, we simulate a range of NC ratios but substituting the LVD system with the BMR system. These results can be seen in Figure 6.8.

In Figure 6.9 we show the relationship between peak concentration of species C and NC ratio.

The arrival time of species C as a function of NC ratio is presented in Figure 6.10.

The trends in Figure 6.8, Figure 6.9 and Figure 6.10

As discussed in chapter 2, concentration measured at one position in a cell does not represent the entire state of the system, specifically concentration gradients which may have a significant effect on behaviour. The simulations in *SS3 results* are repeated, but in this instance the BMR system is used instead of the LVD system. The results of this new set of simulations are shown in Figure 6.11.

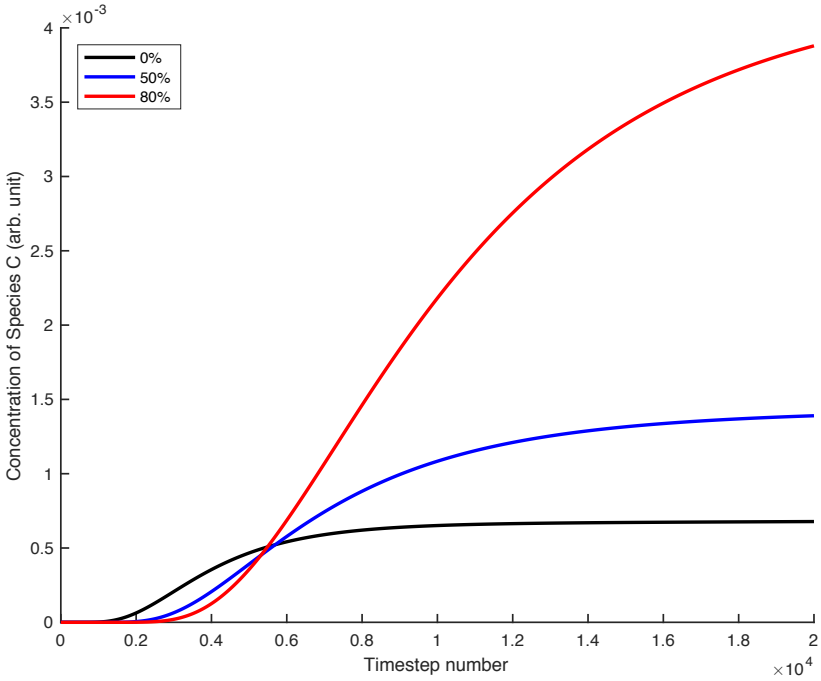


Figure 6.7: *SS4 results*:Time series of species C concentration in the BMR system. As with the LVD system, arrival time increases with NC ratio, however in contrast, peak concentration appears to consistently increase with increasing NC ratio. Finally, in virtual cells with lower NC ratios, steady state behaviour is achieved more rapidly than at greater NC ratios. Percentages refer to NC ratios.

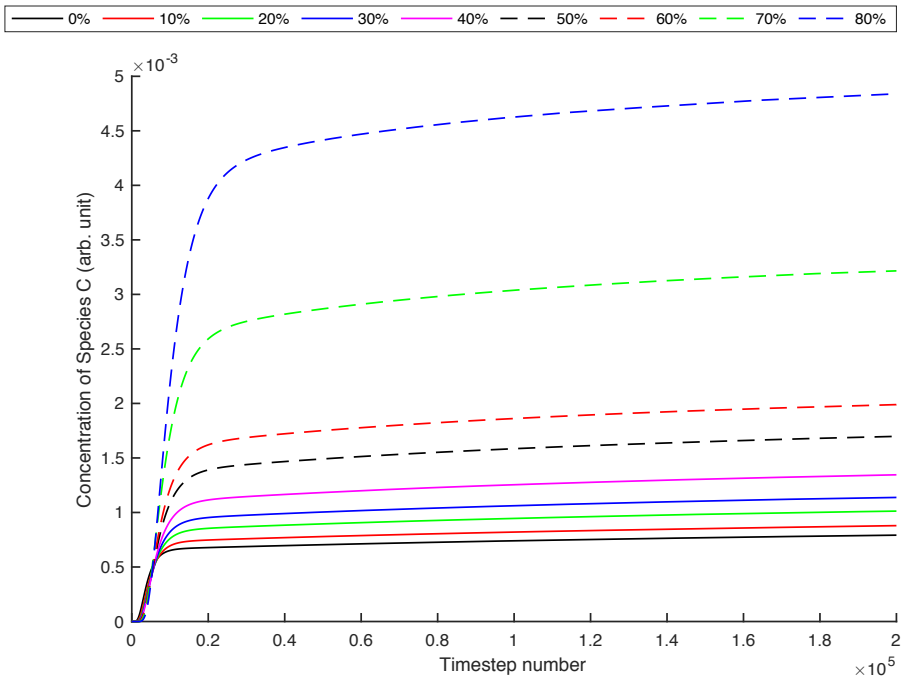


Figure 6.8: *SS5 results*:Time series of species C concentration, in the BMR system, for a range of NC ratios. The relationship here has different qualities to that displayed in Figure 6.3. Increasing NC ratio results in later arrival time, as in the LVD system, but it now also *consistently* results in higher maximum concentrations. Larger NC ratios result in smaller reaction vessels leading to higher concentrations and faster rates of reaction. Percentages in the legend refer to NC ratios.

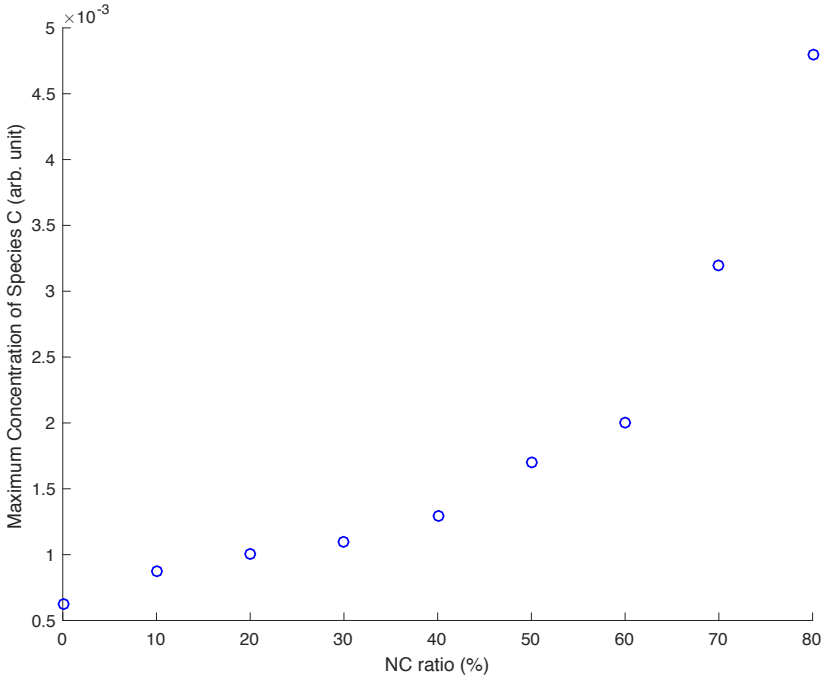


Figure 6.9: *SS5 results*:Peak concentration of species C in cells of different NC ratios. In the BMR system it can be seen that peak concentration increases with NC ratio. This behaviour should be contrasted with the behaviour of the LVD system in Figure 6.4, which displays a different relationship to NC ratio. Peak concentration increases due to increasingly small reaction volumes.

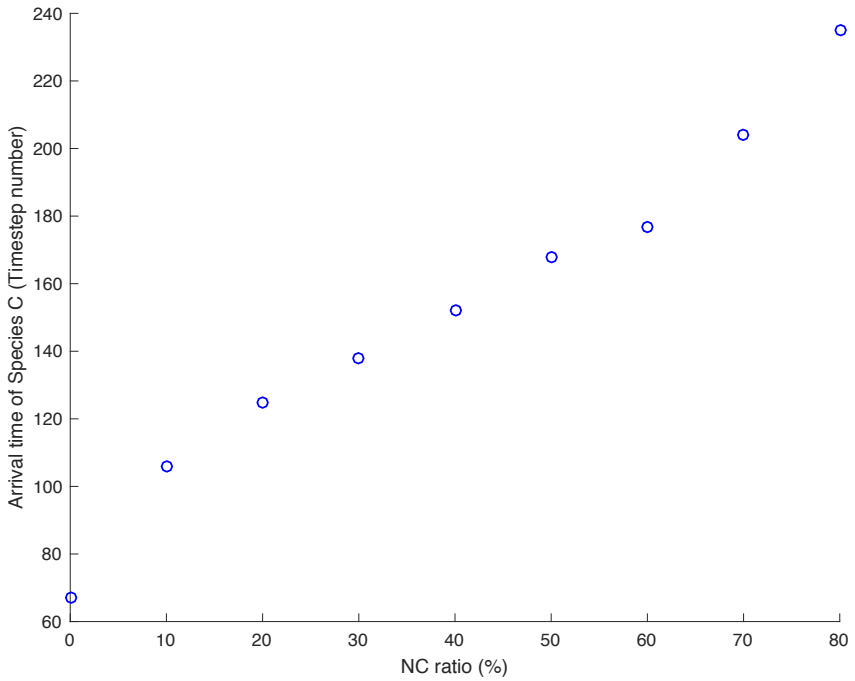
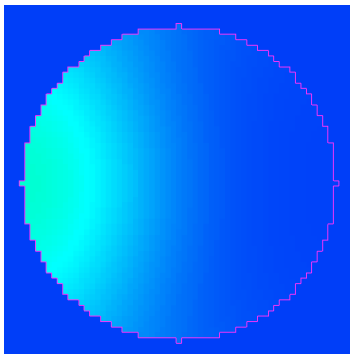
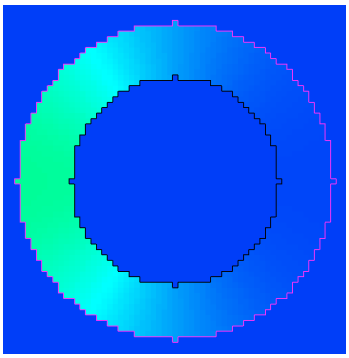


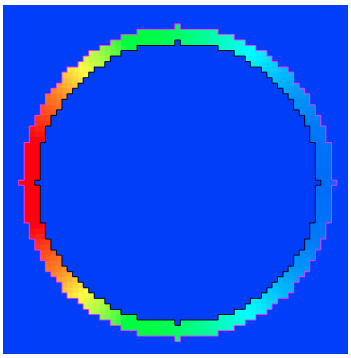
Figure 6.10: *SS 5 results*: Arrival time of species C. The arrival time of species C increases with NC ratio. The relationship between arrival time and NC ratio is similar in both the BMR and the LVD systems (Figure 6.5). Arrival times increase as the chemicals have to navigate increasingly large nuclei inside of the cell.



(a) 0% NC ratio



(b) 10% NC ratio



(c) 50% NC ratio

Figure 6.11: *SS6 results* Concentration gradients of species C in the in the BMR system at different NC ratios. Gradients captured when wavefronts reach the same position. The impact of NC ratio on the state of the system differs considerably depending on the reaction system. Larger NC ratios have a sizeable effect upon the gradient of concentration. This observation is consistent with the graph of peak concentration in Figure 6.9. On the other hand NC ratio seems to have little effect upon the shape of the wavefront. This result should be contrasted with the concentration gradients of the LVD system shown in Figure 6.6. In the LVD system, the concentration gradient appears to be less sensitive to the NC ratio than in the BMR system, but the shape of the wave of concentration is much more dependent upon the NC ratio.

6.2 Nuclear positioning alters virtual pathway behaviour

As discussed in chapter 2, the position of the nucleus can be vital to cell function, with irregularities in position implicated in several diseases. The simulations in this subsection show the link between nuclear position and chemical response.

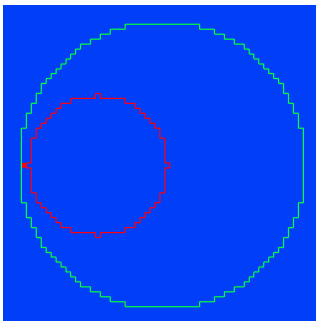
We begin investigating the impact of nuclear position upon virtual pathway behaviour by placing the nucleus in four "extreme" positions within the cell, at 10% and 50% NC ratios, and comparing the concentration gradients and reaction trajectories. These extreme positions at both NC ratios can be seen in Figure 6.12. It is impossible to systematically place the nucleus in all locations within the cell at each NC ratio, therefore we next place the nucleus at random locations within the cell at 0% - 80% in the BMR system. Using these results, we analyse the difference in peak concentration and arrival time between these randomly placed nuclei and the centrally placed nuclei from SS2 and SS5. Finally we place a 10% nucleus in two selected positions within the cell and analyse the concentration gradients produced. These simulations are repeated for the LVD and BMR system.

6.2.1 The LVD system

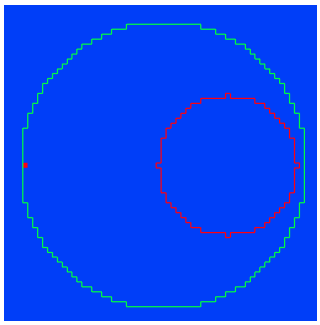
The spatial properties of the LVD system are explored first. In Figure 6.13, the effect of repositioning the nucleus inside the cell is explored by positioning the nucleus at four "extreme positions" at 10% NC ratio.

We then repeat SS7 but in this set of simulations the NC ratio is increased to 50%, representative of an immature blast erythrocyte [Turgeon, 2005]. These results are shown in Figure 6.14.

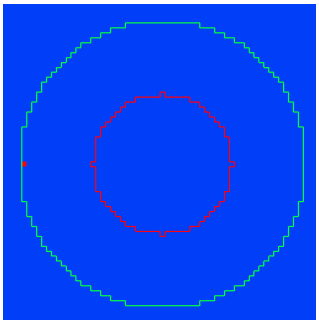
Increasing the NC ratio from 10% (Figure 6.13) to 50% (Figure 6.14) gives nuclear position a much greater impact. At 10% NC ratio, positioning the nucleus far away from both OP and MP (virtual cell 4) has comparatively little effect on the reaction trajectories, but when NC ratio is increased to 50% the



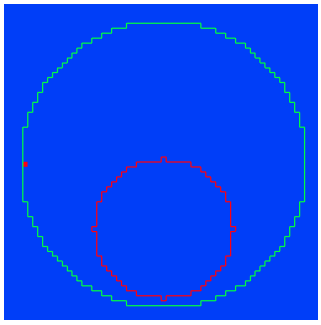
(a) Virtual Cell 1: 10% NC ratio nucleus placed next to OP



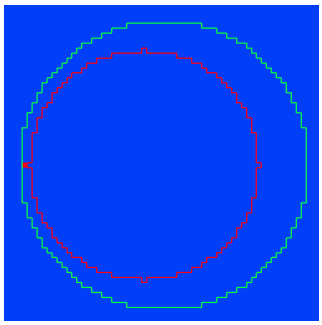
(b) Virtual Cell 2: 10% NC ratio nucleus placed next to MP



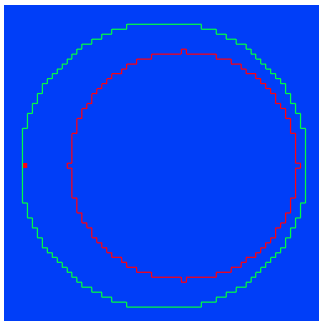
(c) Virtual Cell 3: 10% NC ratio nucleus placed centrally



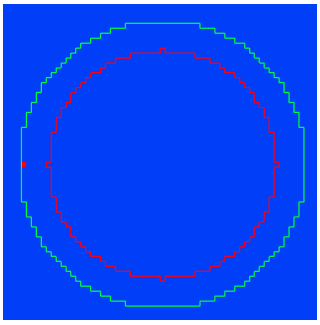
(d) Virtual Cell 4: 10% NC ratio nucleus placed as far away as possible from both OP and MP



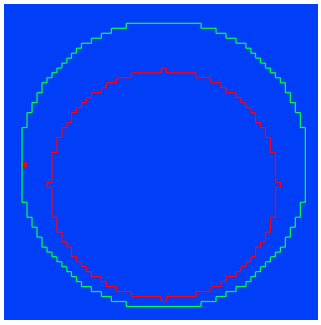
(e) Virtual Cell 5: 50% NC ratio nucleus placed next to OP



(f) Virtual Cell 6: 50% NC ratio nucleus placed next to MP



(g) Virtual Cell 7: 50% NC ratio nucleus placed centrally



(h) Virtual Cell 8: 50% NC ratio nucleus placed as far away as possible from both OP and MP

Figure 6.12: Initial configurations of virtual cells, each with different nuclear positions and NC ratios. Refer to Figure 6.1 for the position of OP and MP.

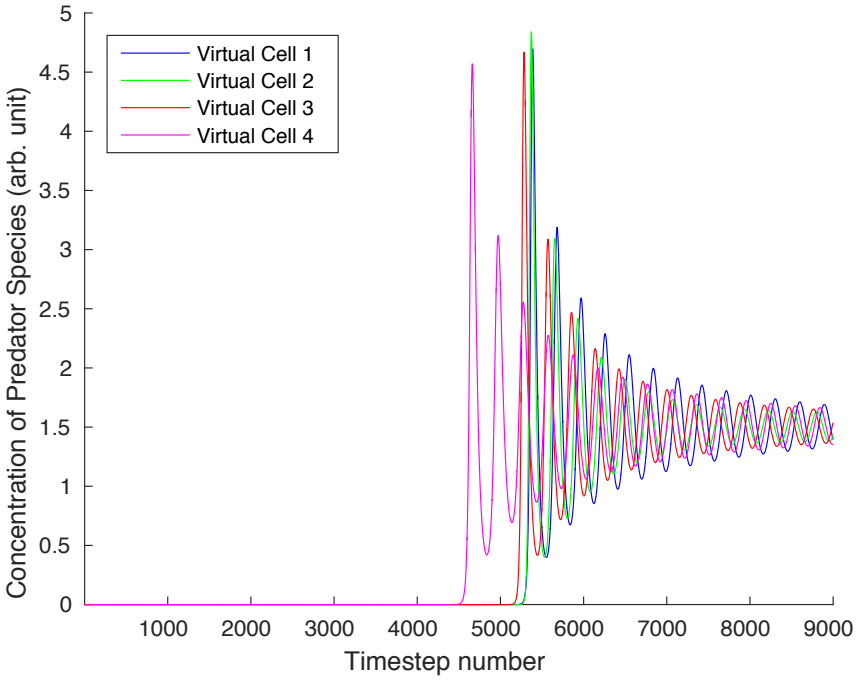


Figure 6.13: *SS7 results*:10% NC ratio virtual cells. Figure show reaction trajectories for four different initial configurations. Virtual cell 1 (nucleus positioned just in front of the OP) shows the latest arrival time, closely followed by virtual cell 2 (Nucleus placed next to MP) and virtual cell 3 (Nucleus placed centrally). Placing the nucleus as far away as possible from both OP and MP results in the earliest arrival time by a large margin. Peak concentration is the highest in virtual cell 2 and lowest in virtual cell 4.

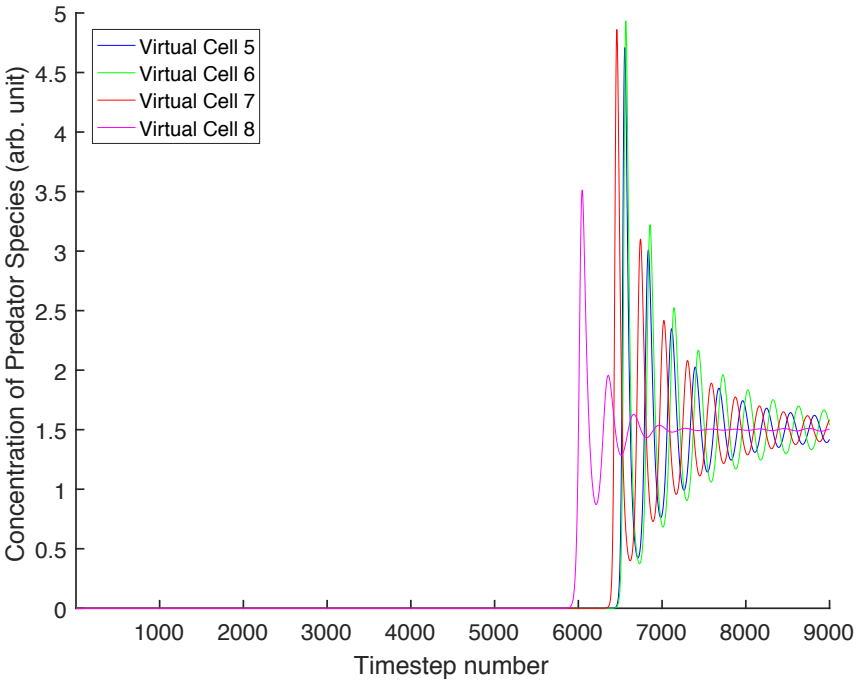


Figure 6.14: *SS8 results*:50% NC ratio virtual cells. Reaction trajectories for four different initial configurations. These results show a large difference in concentration time series between virtual cell 8 and other configurations. The peak concentration of *Predator* in virtual cell 8, with the nucleus located furthest from both OP and MP, is much less than in the other virtual cells. Peak concentration in virtual cell 6 (with its nucleus located next to MP) is the greatest. In addition, the long term behaviour of the *Predator* concentration in virtual cell 8 is completely different to the other cells, exhibiting a highly damped profile.

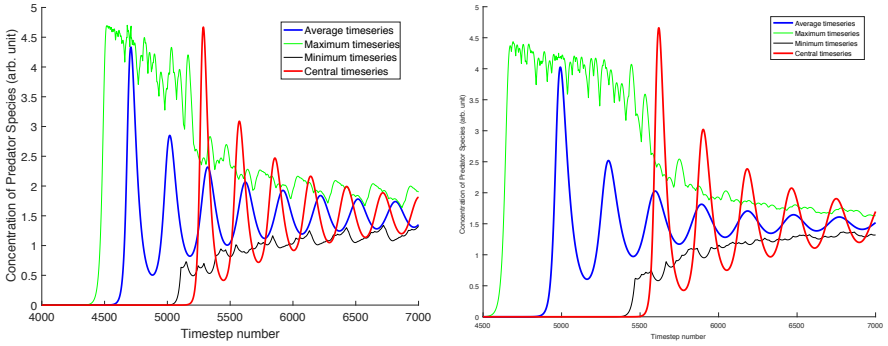
analogous position in virtual cell 8 results in a much reduced peak concentration. Damping behaviour in virtual cells 4 and 8 is also different, at 50% NC ratio the predator concentration timeseries is highly damped. Peak concentration in virtual cells 2 and 6 (with their nuclei located next to MP) is the greatest in both cells. Arrival time behaviour in both virtual cells is consistent.

Whilst exciting, results in SS7 and SS8 only explore a small amount of the state space of experimental setup. By randomly positioning the nucleus over many simulations and NC ratios we can draw a more complete picture of nuclear position dependent behaviour. For each NC ratio and virtual pathway 100 random virtual cells, each with different nuclear positions, are generated. Given that in some healthy cells the nucleus can occupy at least 80% of the total volume of the cell, it is vital to explore nuclear positioning at a variety of NC ratios. With this in mind, the NC ratios explored in this section are 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70% and 80%. For each NC ratio we average the time series produced by the random nuclear positions; we call this time series the mean position time series. We also compute the maximum and minimum values for the concentration of *Predator* species at each point in the time series. We compare the behaviour of the mean position time series with the time series produced by the centrally located nucleus. These results are seen in Figure 6.15. This result shows how varying nuclear position alters the trajectory of pathway behaviour.

Figure 6.16 compares the peak concentration of *Predator* species of the average random time series and the single central nucleus for the LVD system as a function of NC ratio.

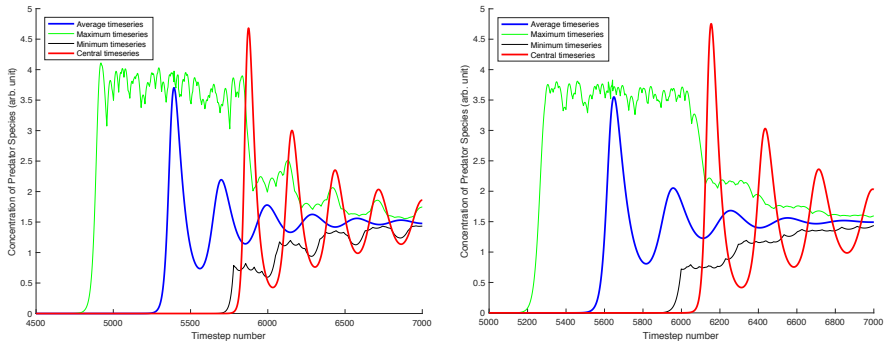
We continue our analysis of the LVD system's response to nuclear positioning by examining the differences in arrival time between centrally located nuclei and randomly located nuclei. Figure 6.17 compares the arrival time of *Predator* species of the average time series and the single central nucleus for the LVD system as a function of NC ratio.

Next we review the sensitivity of concentration gradients to nuclear position



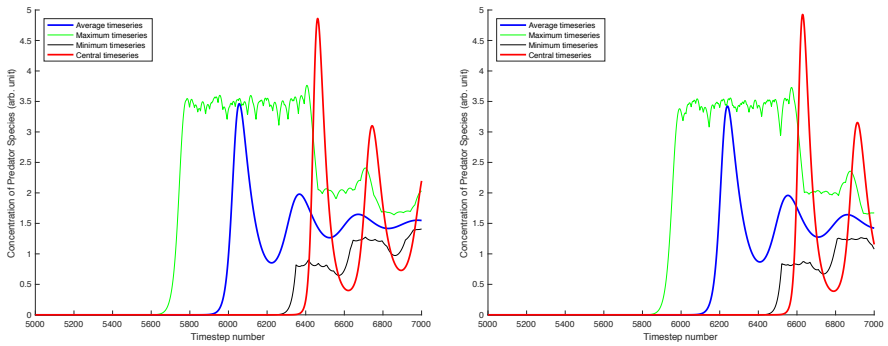
(a) 10% clutter

(b) 20% clutter



(c) 30% clutter

(d) 40% clutter



(e) 50% clutter

(f) 60% clutter

Figure 6.15: *SS9 results*:Graphs of *Predator* species behaviour, showing the effect of nuclear positioning. The mean time series of the random positions, the time series of the central position (from SS2) and the maximum and minimum concentrations at each point are displayed. In these results it can be seen that the centre is the most disruptive position, ensuring the latest arrival time *and* the highest peak concentration in the LVD system. Although the highest concentrations in Figure 6.13 and Figure 6.14 are due to the nucleus being positioned at MP it is unlikely that the nucleus would be randomly positioned at MP. As the NC ratio increases the difference in peak concentration between the centrally positioned nucleus and average nucleus increases.

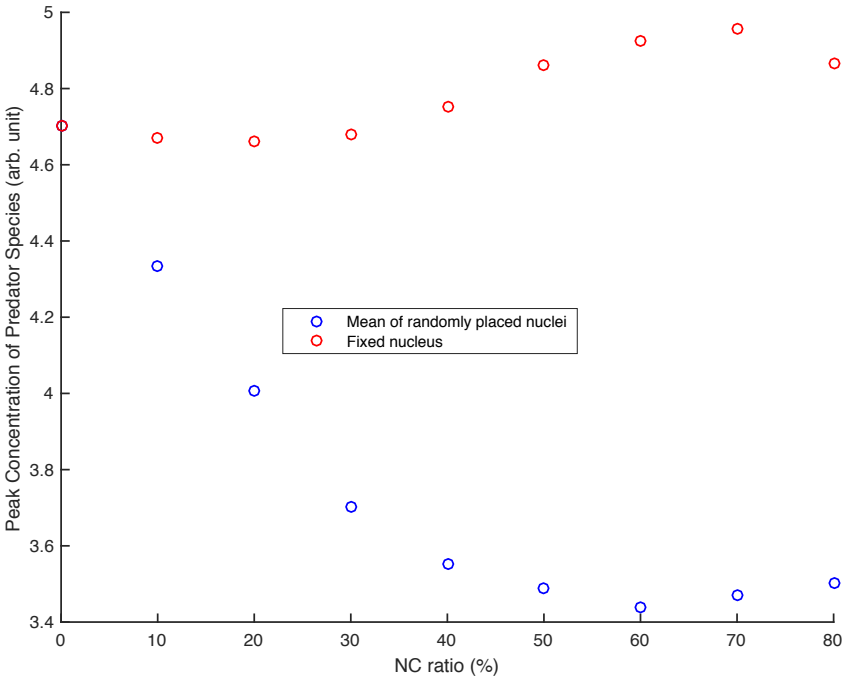


Figure 6.16: *SS9 results*:Peak concentration of *Predator* species in the LVD system. Graph also shows peak concentration from central position timeseries from SS2. This result confirms our initial observations on the importance of central nuclear positioning.

(Figure 6.18) Figure 6.18(a) and Figure 6.18(b) show the difference in wavefront shape when the nucleus is repositioned, and the subtle difference in concentration.

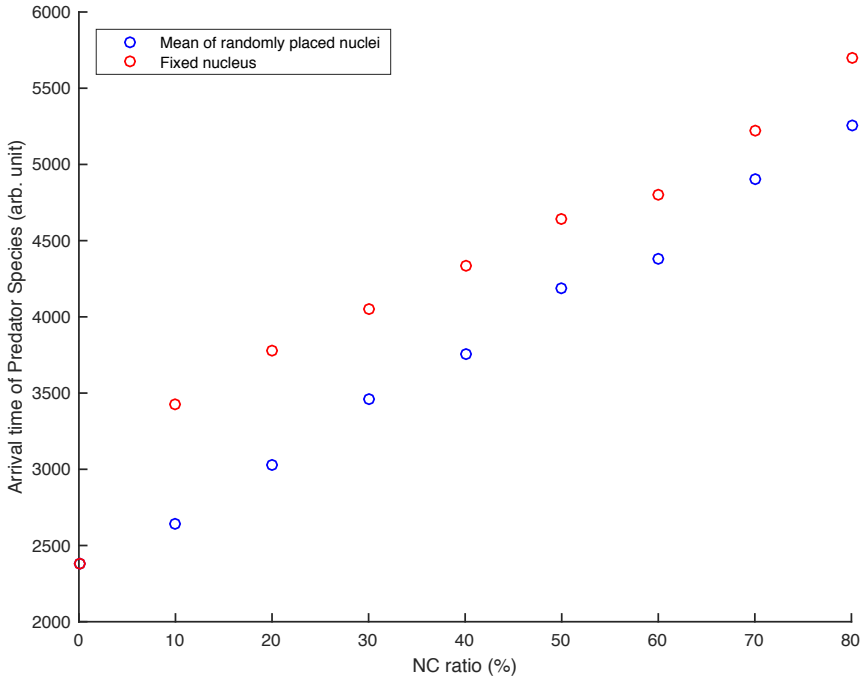
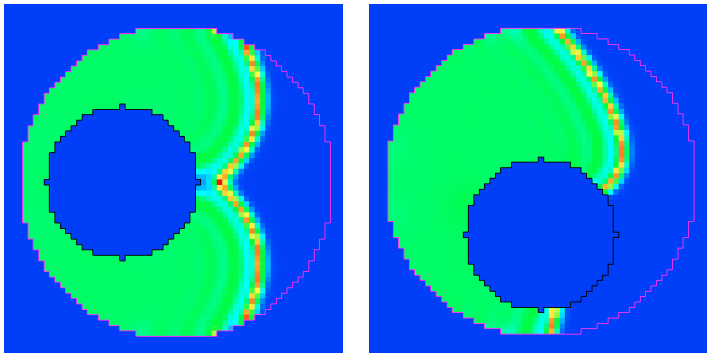


Figure 6.17: *SS9 results*:Arrival time concentration of *Predator* species in the LVD system. Graph also shows arrival time from central position timeseries from SS2. For both centrally positioned nuclei and randomly positioned nuclei arrival time increase approximately linearly. This graph confirms our initial observations on the importance of the central position; centrally located nuclei result in later arrival times than randomly located nuclei at all NC ratios.



(a) 10% in the LVD system (b) 10% in the LVD system

Figure 6.18: *SS10 results*:Wavefronts captured at the same point in time using different nuclear positions, both at 10% NC ratio. Wavefront shape, concentration gradient and concentration all change when the nucleus is moved.

6.2.2 The BMR system

Having presented the LVD system's spatial sensitivity to nuclear positioning, we now examine the response of the BMR system to nuclear positioning. We repeat SS7 and SS8, instead using the BMR system, to form an initial idea of this system's spatial sensitivity. SS11 (Figure 6.19) demonstrates nuclear position sensitivity, in virtual cells 1-4 (Figure 6.12) and SS12 (Figure 6.20) demonstrates nuclear position sensitivity, in virtual cells 5-8.

In Figure 6.20, the NC ratio is increased to 50%.

Once again, we explore more of the initial condition state space through random positioning. The results of this set of simulations can be seen in Figure 6.21. This set of simulations is a repeat of those in SS9, but with the LVD system replaced with the BMR system.

Next we repeat the peak concentration analysis that was performed for the LVD system. Figure 6.22 compares the peak concentration of species C for the average random time series and the single central nucleus for the BMR system as a function of NC ratio.

Figure 6.23 shows how arrival time of species C differs between centrally located nuclei, and the average randomly located nuclei in the BMR system.

In Figure 6.24, we evaluate the BMR system's concentration gradient sensitivity to nuclear position.

We finalise our study of the effect of nuclear positioning on arrival time, by comparing the standard deviation of arrival times at different NC ratios in both of the reaction systems. These results can be seen in Figure 6.25.

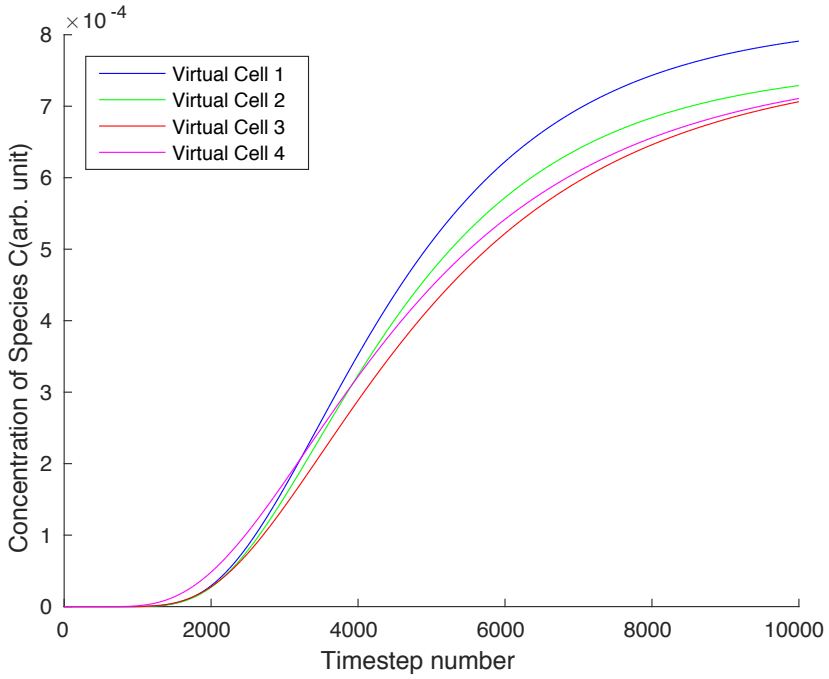


Figure 6.19: *SS11 results*:10% NC ratio virtual cells. Figure shows reaction trajectories for four different initial configurations. Virtual cell 4 has the earliest arrival time. Virtual cell 1 displays the greatest peak concentration and virtual cells 3 and 4 show the lowest peak concentration. These results stand in contrast with Figure 6.13, where virtual cell 2 displays the greatest concentration and virtual cell 1 the smallest value.

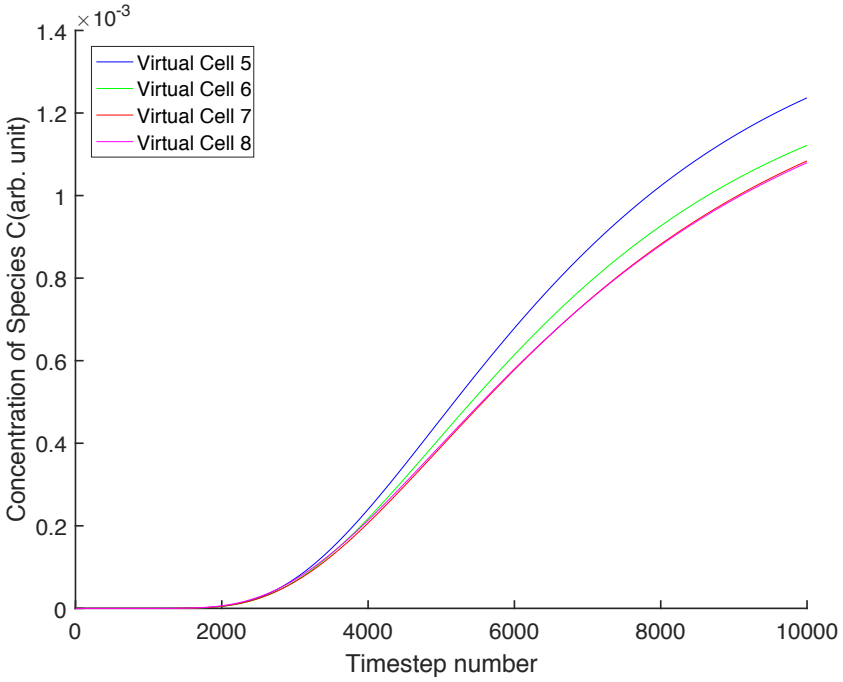
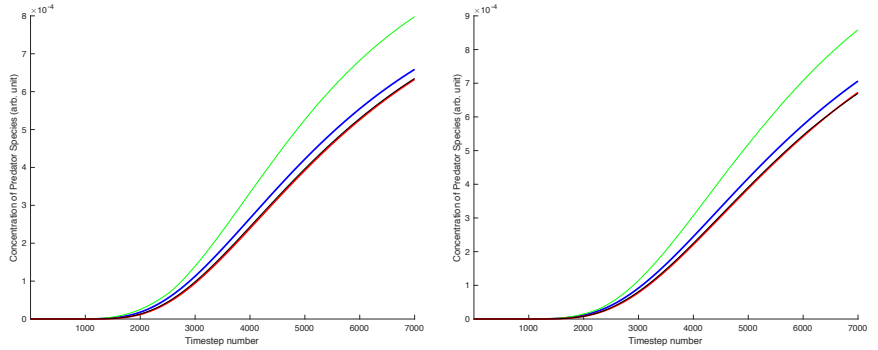
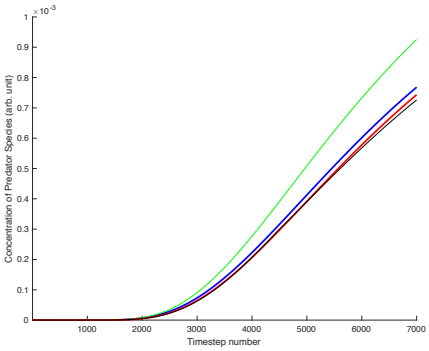


Figure 6.20: *SS12 results*:50% NC ratio virtual cells. Figure show reaction trajectories for four different initial configurations. Virtual cells 1 and 2 still occupy the first and second positions for peak concentration respectively, virtual cells 3 and 4 have very similar concentration trajectories. The difference in peak concentration (between the time series displaying the highest and lowest concentrations) is greater in the set of virtual cells with a higher NC ratio (SS12) than in the set with the lower NC ratio (SS11). This difference in peak concentration (between 10% and 50%) is much smaller in the BMR system compared to the LVD system.

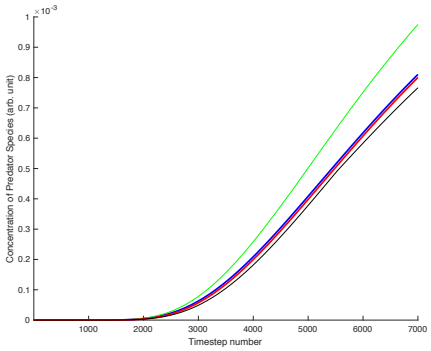


(a) 30% clutter

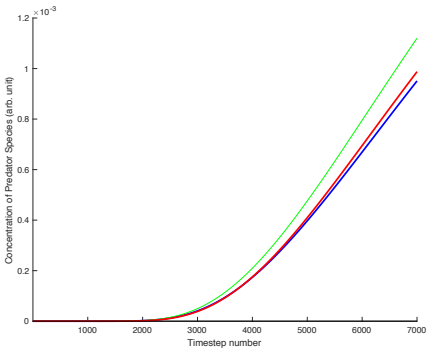
(b) 40% clutter



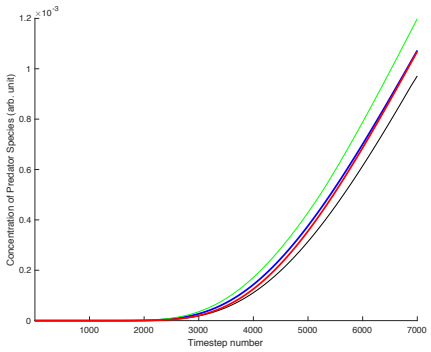
(c) 50% clutter



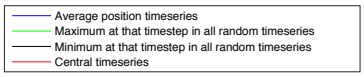
(d) 60% clutter



(e) 70% clutter



(f) 80% clutter



(g) legend

Figure 6.21: *SS13 results*:Graphs of species C concentration in the BMR system, showing the difference in random and central positioning. The mean time series of the random positions, the time series of the central position (from SS5) and the maximum and minimum concentrations at each point are displayed. At all NC ratios in this SS, the central position does not result in the highest concentration, as in Figure 6.19 and Figure 6.20. Instead, as NC ratio increases, the peak concentration in the central nucleus simulation approaches the maximum of the random positions. The behaviour in this SS should be contrasted with SS9 in Figure 6.15.

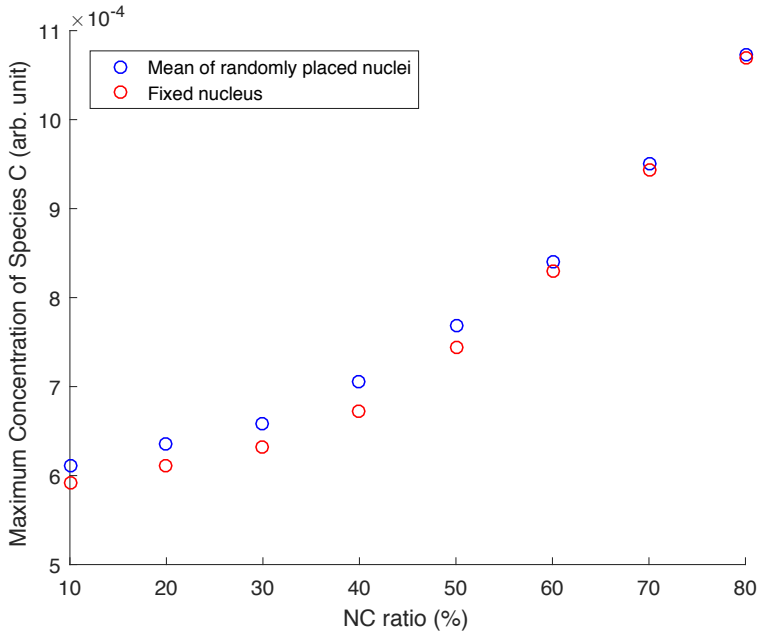


Figure 6.22: *SS13 results*:Peak concentration of species C in the BMR system. Graph also shows peak concentration from central position timeseries from SS5. This result shows an insignificant difference between peak concentration in simulations with fixed and average randomly positioned nuclei. This similarity in concentration between fixed and random nuclei should be contrasted with the analogous result for the LVD system shown Figure 6.16, which is significant.

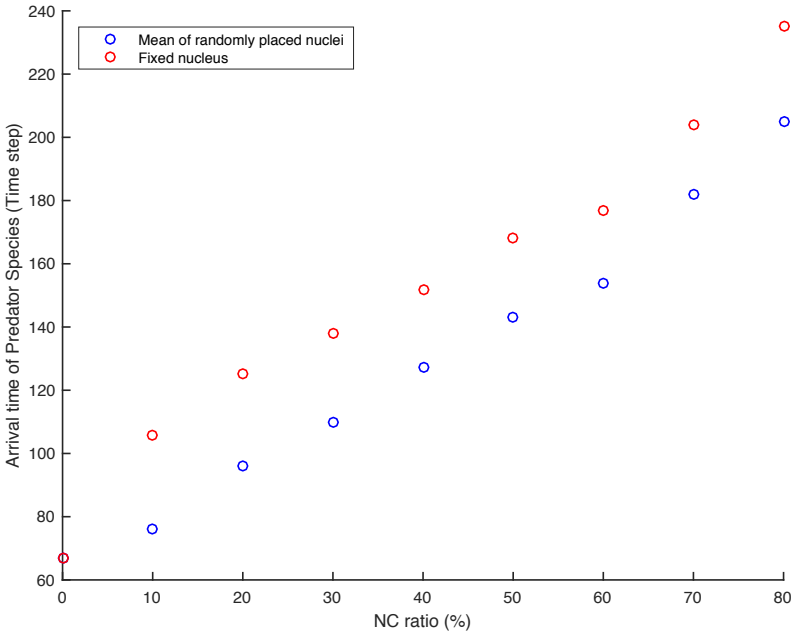
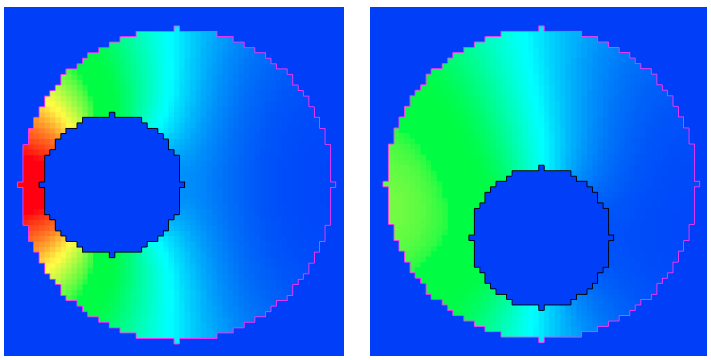


Figure 6.23: *SS13 results*:Arrival time of species C in the BMR system. Graph also shows arrival time from central position timeseries from SS5. In this result, the fixed nucleus displays the later arrival time. The results in Figure 6.17 show similar behaviour. These results indicate the independence of arrival time behaviour from the type of reaction system and the dependence of arrival time behaviour on both NC ratio and nuclear position.



(a) 10% in the BMR system (b) 10% in the BMR system

Figure 6.24: *SS14 results*:Wavefronts captured at the same point in space for different values of NC. As in Figure 6.11, a clear difference in concentration gradient is observed when the nucleus is repositioned within the cell. In addition, we also observe a subtle shift in wavefront shape. This behaviour is unlike that displayed within the LVD system (Figure 6.18) where wavefront shape was dramatically affected but there was less of a difference in concentration gradient.

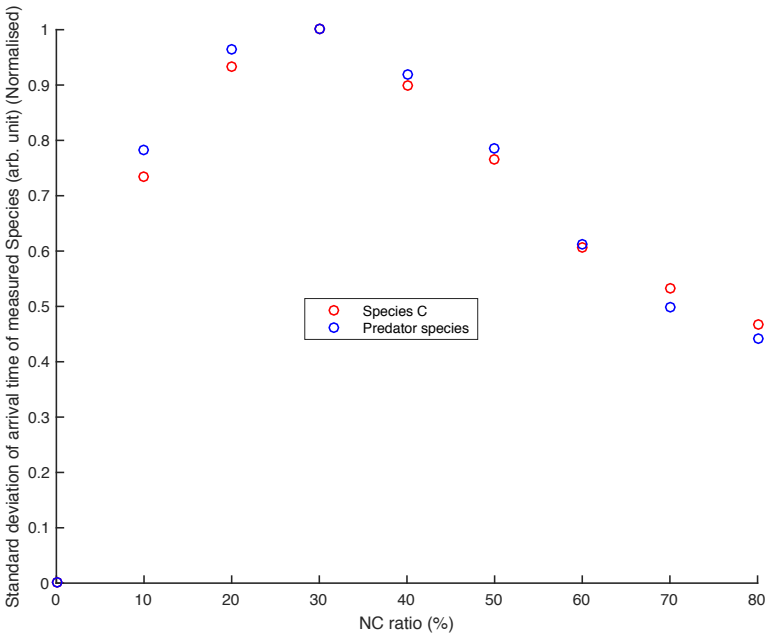


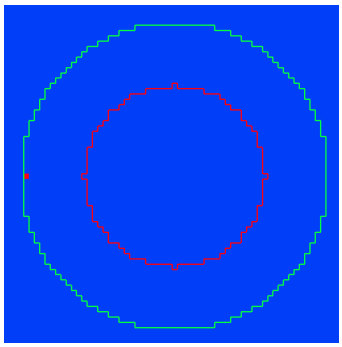
Figure 6.25: Normalised standard deviation of arrival times in both the LVD and BMR system from SS9 and SS13. In both systems, standard deviation increases until it peaks at 30% NC ratio and then gradually decreases. Figure 6.25 shows very similar behaviour for both reaction-diffusion systems. This result provides further evidence of the invariance of arrival time behaviour to the reaction system.

6.3 Distribution of internal cellular clutter alters the behaviour of virtual pathways

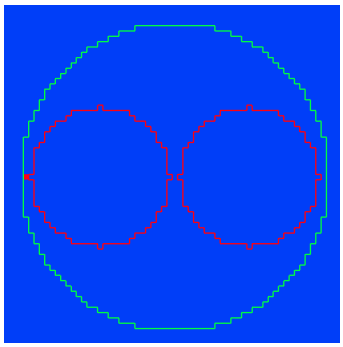
In Subsection 2.5.3 it was demonstrated that internal cellular volume can be distributed in a range of configurations; concentrated all in one piece or sub divided into pieces of varying size. In this section we simulate the impact that the distribution of internal volume has on the regulation of model cellular pathways.

Internal volume is distributed by taking an SC ratio, splitting the volume into a number of "subvolumes", and then randomly positioning them within the virtual cell. The number of subdivisions varies depending upon the SC ratio. Up to and including SC ratios of 20%, the subdivisions are of size 5% (these 5% subdivisions are located randomly within the cell). For example in a cell with a 15% SC ratio three randomly positioned 5% subdivisions would be placed within the cell. Beyond 20% SC ratio it becomes mathematically impossible to place enough sub-divisions of equal volume into the virtual cell. Therefore, in cells with SC ratios greater than 20%, we have to introduce sub-divisions of heterogeneous sizes, mimicking organelle size differences in real cells. Past 50%, it becomes challenging to place spherical virtual organelles in the cell in any configuration.

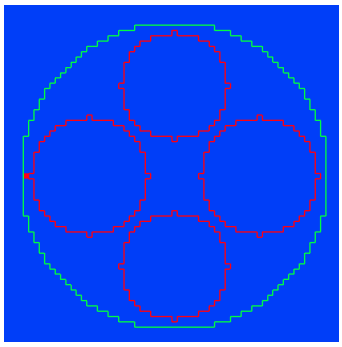
We begin investigating the impact of nuclear position upon virtual pathways by using the virtual cells in Figure 6.26 and comparing the produced concentration gradients and reaction trajectories. It is impossible to systematically generate all substructure divisions and positions at each SC ratio, therefore we randomly generate substructure configurations within the cell at 0% - 50%. Using these results, we analyse the difference in peak concentration between randomly generated configurations, randomly placed nuclei (from SS9 and SS13) and centrally placed nuclei from SS2 and SS5.



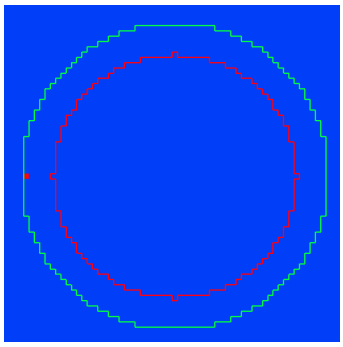
(a) Virtual Cell 9: 20% SC ratio containing one substructure, centrally located



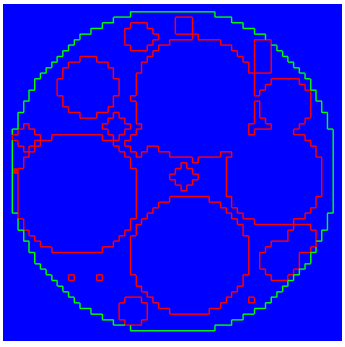
(b) Virtual Cell 10: 20% SC ratio containing two substructures, of equivalent total volume to virtual cell 9



(c) Virtual Cell 11: 20% SC ratio containing four substructures, of equivalent total volume to virtual cell 9



(d) Virtual Cell 12: 50% SC ratio containing one substructure, centrally located. This virtual cell is identical to 7, but it is reproduced here for convenience.



(e) Virtual Cell 13: 50% SC ratio containing many substructures

Figure 6.26: Figure shows initial configurations, each with different substructure distributions at 20% and 50% SC ratios

6.3.1 The LVD system

We begin by exploring the behaviour of the LVD system. Figure 6.27 shows the results from deterministically exploring substructure configuration.

Figure 6.28 compares the concentration trajectories of virtual cells 12 and 13 (one with a centrally located nucleus and another with an equivalent volume of substructure randomly distributed throughout the cell).

The difference in timeseries behaviour between centrally located and distributed substructures is determined by SC ratio. At 20% SC ratio (Figure 6.27) the difference between the two is much less pronounced than at 50% NC ratio (Figure 6.28)

We then explore more of the state space through randomisation. For each SC ratio, 100 random virtual cells are generated, each with their subvolumes in a different position. In addition to this, for each SC ratio, 100 random virtual cells are generated, each with their nucleus in a different position (as in Subsection 2.5.3). Finally, for each SC ratio, a virtual cell with a single centrally positioned nucleus is generated (as in Section 6.1). These three sets of measurements that comprise this subsection allow us to contrast the behaviour of centrally positioned nuclei, randomly positioned nuclei and randomly distributed sub-divisions of structure. Figure 6.29 shows the time series behaviour of centrally located nuclei, the average of the repositioned nuclei and distributed subvolumes of equivalent volumes.

Figure 6.30 shows peak concentrations in the central, random and distributed spatial configurations at different NC ratios.

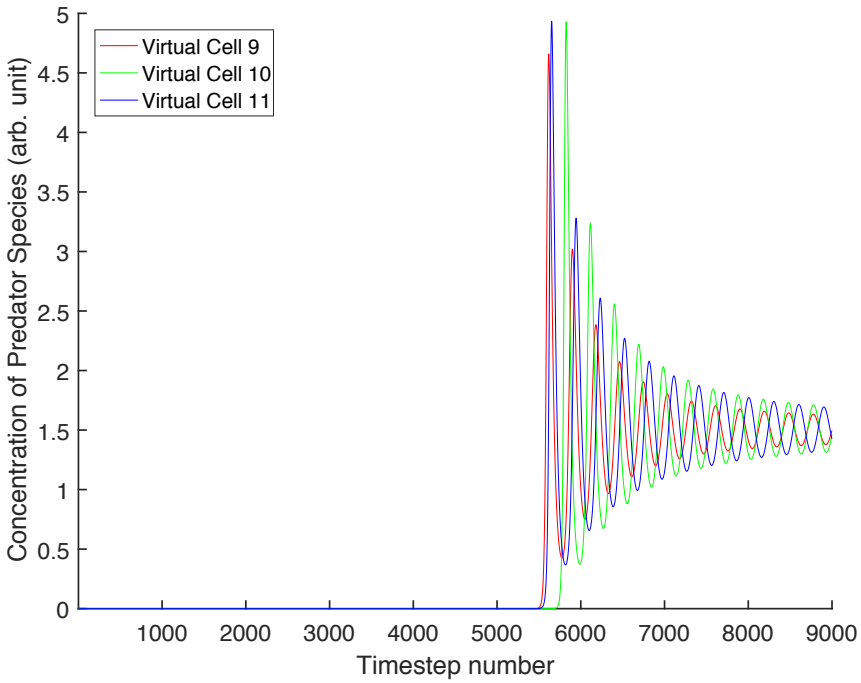


Figure 6.27: *SS15 results*:20% SC ratio virtual cells. Figure shows reaction trajectories for three different initial configurations. Virtual cell 9 gives the lowest peak concentration of *Predator* species, with virtual cells 10 and 11 tied for the greatest peak concentration. Virtual cells 9 and 10 display the earliest arrival time.

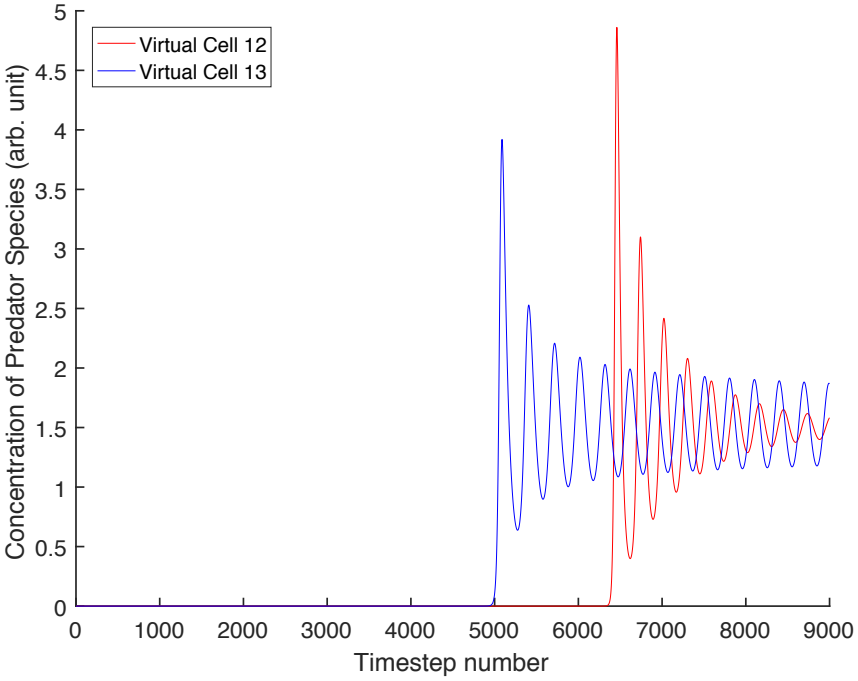
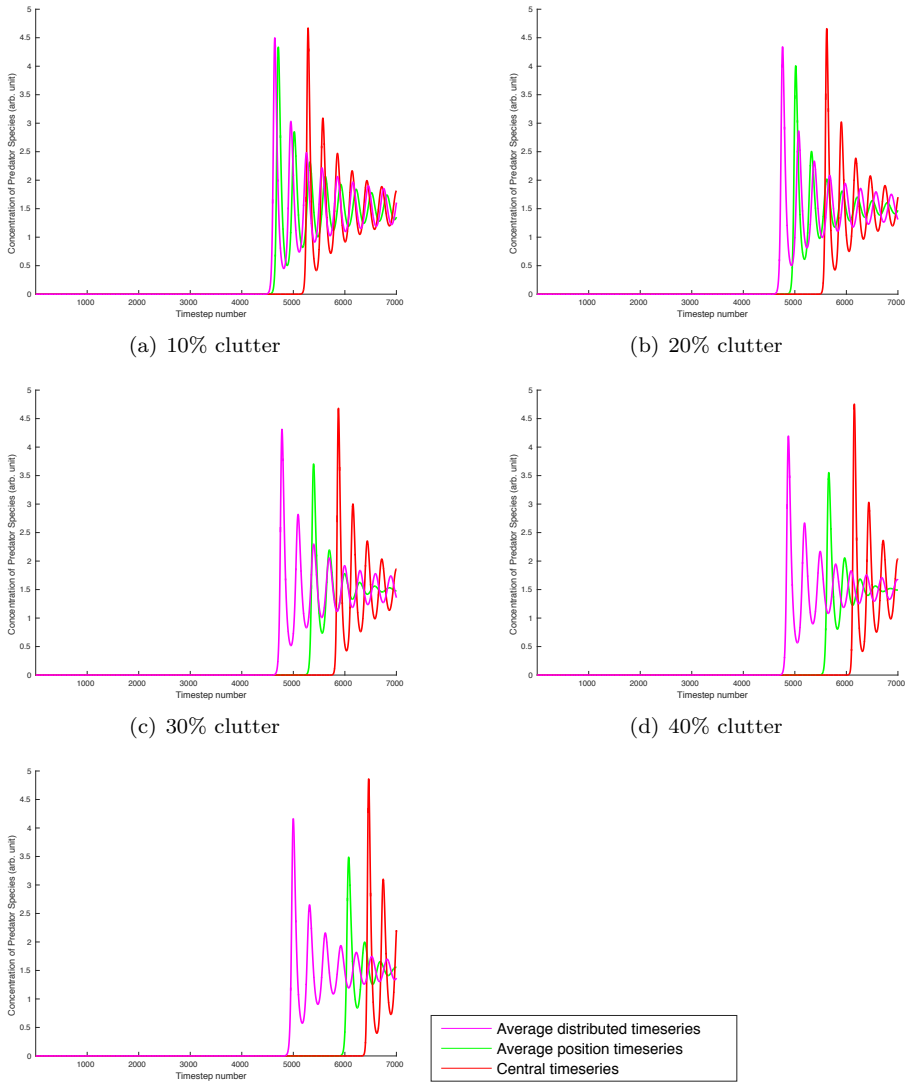


Figure 6.28: *SS16 results*:50% SC ratio virtual cells. Figure shows reaction trajectories for two different initial configurations. Virtual cell 12 (with a single centrally located substructure) results in a much later arrival time than virtual cell 13. Damping behaviour is different in both virtual cells. The amplitude of virtual cell 12 decays much more quickly than virtual cell 12. In addition, the peak concentration in virtual cell 12 is greater than in virtual cell 13.



(a) 10% clutter (b) 20% clutter

(c) 30% clutter (d) 40% clutter

(e) 50% clutter (f) legend

Figure 6.29: *SS17 results*:Graphs of *Predator* species concentration, demonstrating the effect of internal volume distribution. The mean time series of the random positions (SS9), the time series of the central position (SS2) and the mean time series of the randomly distributed sub-volumes are displayed. All quantities are measured in the destination sub-volume. Peak concentration, damping behaviour and arrival time (judged by when the concentration becomes non-zero) are all altered by spatial configuration. In all results, systems with distributed volume result in the earliest arrival time, beating randomly positioned nuclei and centrally positioned nuclei. Additionally, in all results the time series that represents the distributed volume has a peak concentration in between that of the centrally and randomly positioned nuclei. The greater the NC ratio, the more pronounced the effects become. It is obvious that as NC ratio increases, the difference in arrival time of species within the distributed environments increases. The same appears to be true for the relationship between peak concentration and NC ratio, although it appears to be more subtle.

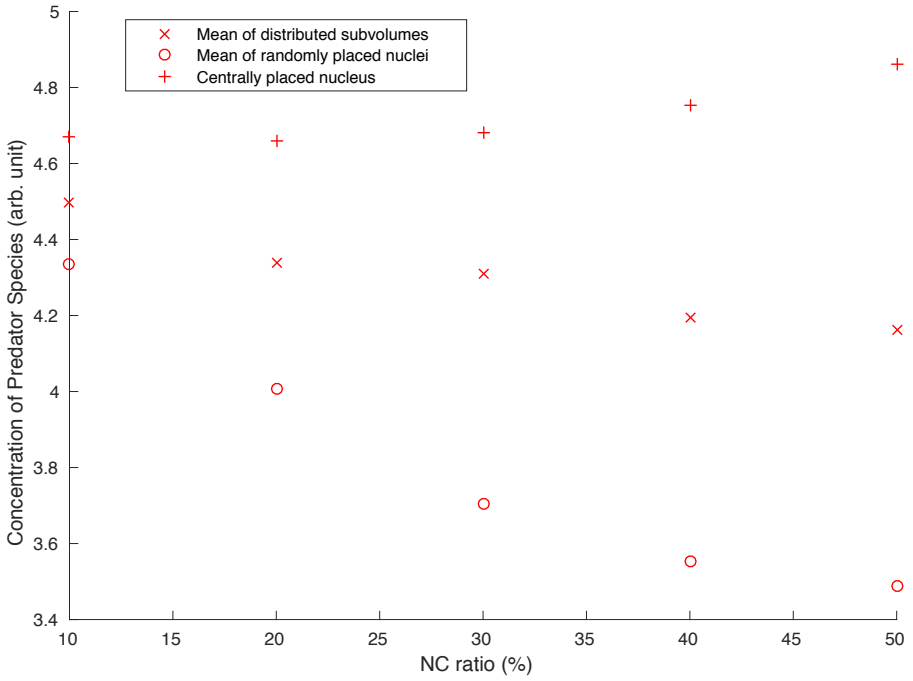


Figure 6.30: *SS17 results*:Peak concentration of *Predator* species in the LVD system. Peak concentrations measured in virtual cells with centrally located nuclei (SS2), randomly located nuclei (SS9) and randomly distributed structures. Observations in this result are consistent with the behaviours in Figure 6.29; distributed concentrations occupy an intermediate positions between randomly and centrally located nuclei.

6.3.2 The BMR system

Having examined the response of the LVD system to substructure configuration, we turn our attention to the BMR system. We repeat SS15 and SS16 substituting the LVD system for the BMR system. Figure 6.31 shows a systematic progression of substructure division at 20% NC ratio. The centrally located single nucleus (Virtual Cell 9) produces the lowest peak concentration, whereas the cells with divided substructures (virtual cells 10 and 11) compete for the greatest peak concentration. This pattern of concentration behaviour is identical to that displayed in the LVD system (Figure 6.31).

The NC ratio is then increased to 50%. Virtual cell 12, with a single centrally located substructure, is compared to virtual cell 13, with randomly distributed substructures. These results are seen in Figure 6.32.

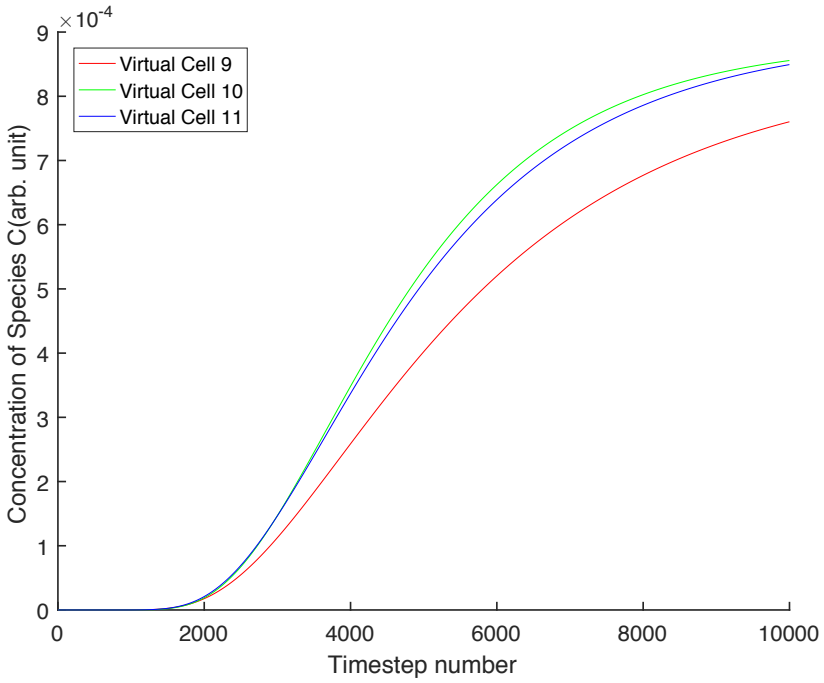


Figure 6.31: *SS18 results*:20% SC ratio virtual cells. Figure shows reaction trajectories for three different initial configurations

The differences in behaviour between the 50% SC ratio cell (Figure 6.32) and the 20% SC ratio cell (Figure 6.31) in the BMR system are remarkably similar to those in the LVD system (Figure 6.27 and Figure 6.28). In both reaction systems, the time series from the distributed subvolume virtual cells resulted in lower peak concentrations, earlier arrival times and earlier equilibrium of concentration.

Following the systematic exploration of the BMR's response to spatial distribution, we explore more of the initial configuration state space by randomly positioning substructures within the cell. The results of these experiments are seen in Figure 6.33.

Figure 6.34 shows the differences in peak concentration of species C in virtual cells with centrally positioned nuclei, randomly positioned nuclei and those with distributed substructure.

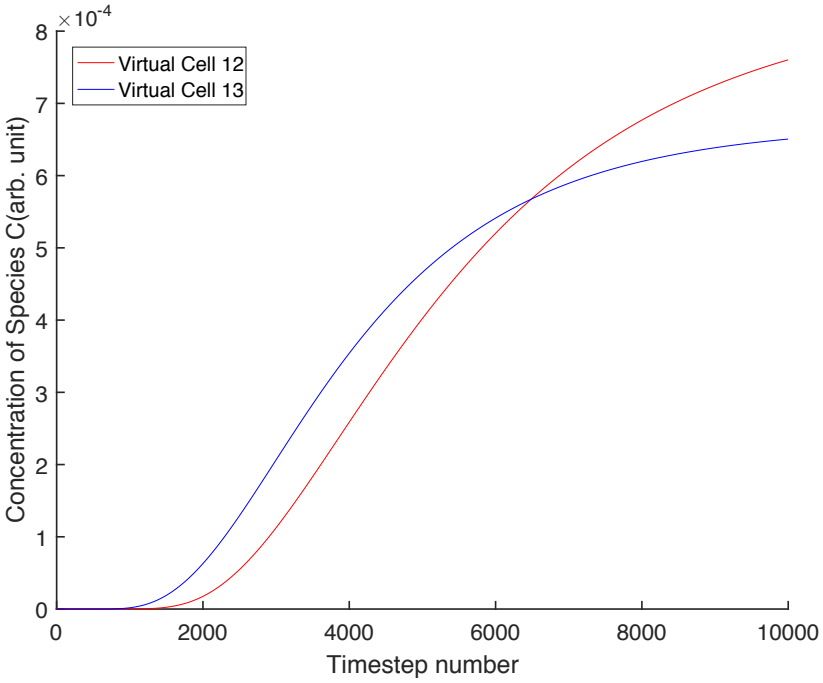
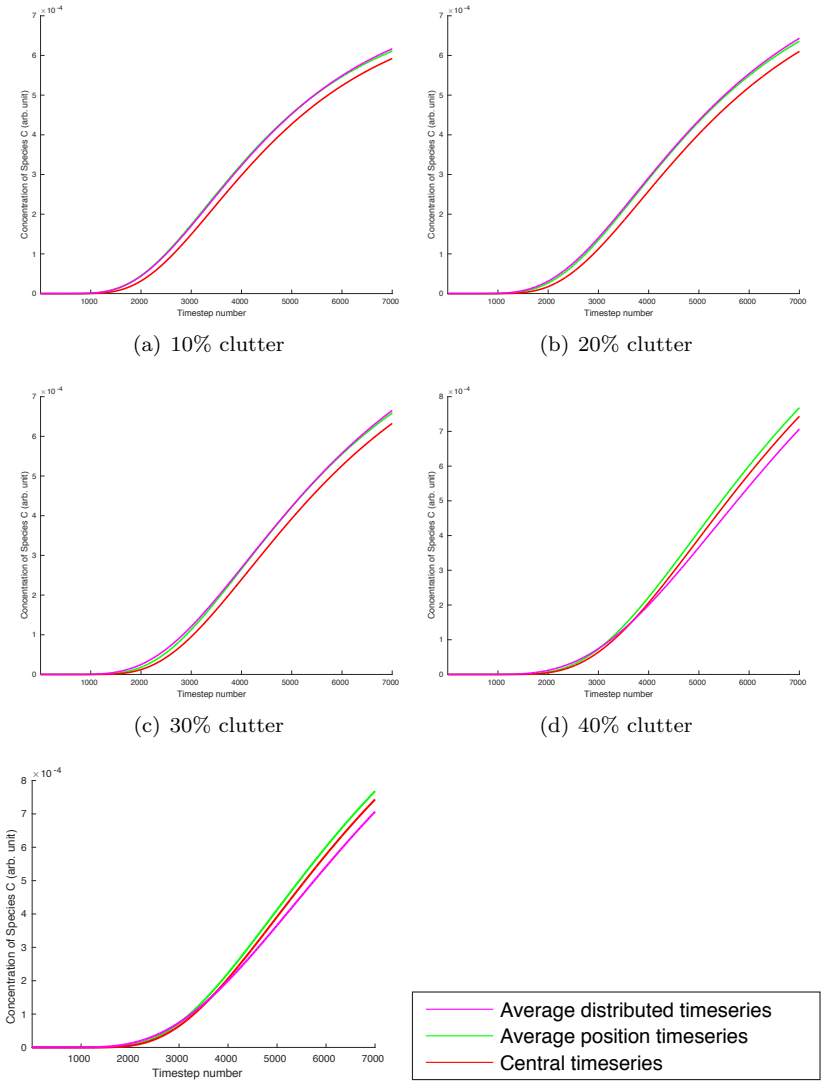


Figure 6.32: *SS19 results*:50% SC ratio virtual cells. Figure shows reaction trajectories for three different initial configurations. This result shows the centrally located nucleus causes a later arrival time and a greater peak concentration.



(a) 10% clutter (b) 20% clutter

(c) 30% clutter (d) 40% clutter

(e) 50% clutter (f) legend

Figure 6.33: *SS20 results*:Graphs of species C concentration, demonstrating the effect of internal volume distribution. The mean time series of the random nuclear positions (SS13), the time series of the central position (SS5) and the mean time series of the randomly distributed sub-volumes are displayed. Once again, the random exploration of initial simulation condition validates the systematic exploration of initial simulation condition. At lower SC ratios, a single centrally positioned subvolume displays lower peak concentration than distributed subvolume (Figure 6.33(a),Figure 6.33(b),Figure 6.33(c)). But once 30% SC ratio is passed, centrally positioned nuclei result in higher peak concentrations. The average position timeseries always results in the greatest concentration.

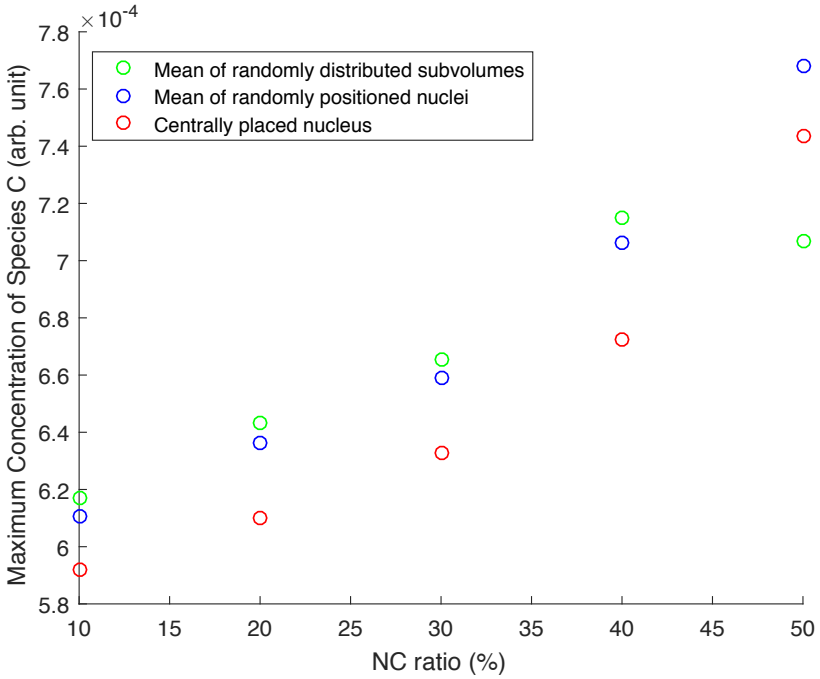


Figure 6.34: *SS21 results*:Peak concentration of species C in the BMR system. Peak concentrations measured in virtual cells with centrally located nuclei (SS5), randomly located nuclei (SS13) and randomly distributed structures. In this results, it is seen that the centrally positioned nucleus never results in the highest peak concentration. Randomly positioned nuclei are nearly indistinguishable from the randomly distributed subvolumes, until an SC ratio of 40% is surpassed. This result should be contrasted with those in Figure 6.30. In Figure 6.30, centrally positioned nuclei display the greatest peak concentration whereas in the BMR system they display the lowest peak concentration, up to and including 40% (Figure 6.34). In addition, the differences in peak concentration between the three different types of virtual cell are much more apparent in the LVD system than in the BMR system.

6.3.3 Analysis of Simulations

Our aim in this work has been to clarify the impact of three specific variables of spatial configuration (NC ratio, nuclear position and structure distribution) through simulation. These aspects of spatial organisation are explored by placing impenetrable spheres, representing nuclei and other organelles, inside a virtual *Coccus* and then permitting a reaction-diffusion system modelling a biochemical pathway to progress. The chemical species involved in the reaction-diffusion systems are placed at the OP (origin point) at the start of the simu-

lation and the chemical concentrations are measured at the MP (measurement point) at the other side of the cell over time to form reaction trajectories. In addition to reaction trajectories we have also measured arrival time and taken snapshots of chemical gradients.

In all our results in this work, it is clear that the greater the proportion of cell volume occupied by structure (nuclei or organelles), the later the arrival time of all chemical species at MP. This result is independent of how the structure is positioned - greater occupancy results in increasing arrival times. Centrally positioned nuclei result in the latest arrival times, whether the pathway was modelled as an LVD or BMR system.

Beyond the effect on arrival time, other results depend on the chosen reaction-diffusion system. Increasing NC ratio in the BMR and LVD systems have different effects on reaction trajectory and peak concentration as shown in Figure 6.4 and Figure 6.9. While both sets of results show a trend, the particular trend is different in each case.

Repositioning the nucleus within the cell also elicits different responses in the different systems (Figure 6.15 Figure 6.21). In the LVD system, no randomly positioned nucleus ever results in a higher peak concentration than nuclei positioned at OP, MP or centrally (Figure 6.13, Figure 6.14, Figure 6.15). The difference in concentration between the highest randomly positioned nucleus and the centrally positioned nucleus increases as the NC ratio increases. The response to repositioning the nucleus in the BMR system is different. In none of the experiments did the central nucleus achieve a peak concentration greater than the highest randomly positioned nucleus. However, as the NC ratio increases, the difference between the two decreases (Figure 6.19, Figure 6.20, Figure 6.21).

In both the BMR and LVD systems, the position of the nucleus has the same effect on the distribution of arrival times with the results in Figure 6.25 showing similar behaviour in both systems, despite possessing different parameters. It appears that the greatest range of arrival times peak at 30% NC.

When the average peak concentrations from the randomly positioned nuclei are plotted with the peak concentration for the centrally positioned nucleus (Figure 6.16 and Figure 6.22), we are able to highlight the effect that changing the same spatial variable has on different reaction-diffusion systems. In the **bmd!** (**bmd!**) system, trends of peak concentration in virtual cells with centrally and randomly positioned nuclei show some similarity. However, in the simulations that use the LVD system, the peak concentrations for randomly and centrally positioned nuclei follow completely different trends.

The greater differences in the LVD system suggest that the LVD system is more sensitive to nuclear position that its BMR counterpart. This statement is borne out by the differences in concentration gradient and wavefront shape when the nucleus is repositioned. In Figure 6.18 and Figure 6.24 we see the effects of nuclear positioning in both systems. In the BMR system the wavefront shapes and concentration gradients are less dramatically affected by repositioning than in the LVD system. In the LVD system, wavefront shapes are shown to be much more sensitive to the positioning of the nucleus.

We then study the effect of distributing internal volume which we describe with the structure to cytoplasm ratio (SC ratio). Again we observe that the results of changing this spatial property depend on the choice of reaction-diffusion system Figure 6.30 and Figure 6.34. Once more it is apparent that the LVD system is much more sensitive to the spatial properties of the cell than the BMR system. The differences in reaction trajectory between randomly positioned substructure and randomly positioned nuclei depend on the choice of reaction diffusion system. In the BMR system, between 10% and 40% occupancy, the difference between the two trajectories are statistically insignificant but at 50% a seemingly non-continuous jump occurs and the difference between the two become significant. In contrast, in the LVD system, the difference between random nucleus position and random substructure distribution is always significant. Furthermore the discrepancy between the two changes in a continuous manner rather than the discontinuous behaviour we observe between 40% and 50% in

the BMR system.

In simulations of the LVD system, peak concentration for the distributed structure is always greater than the peak concentration for the randomly positioned nucleus and always less than the peak concentration for the centrally positioned nucleus. In addition, chemical species in the cells with distributed clutter always arrive at MP first. On average, cells with a non-centrally located nucleus and those who have their structure distributed have much less of an impact on arrival time and in the case of the LVD system, concentration.

Aside from similar arrival time behaviour the BMR system reacts to internal reorganisation in a completely different way to the LVD system. Up to and including 40% occupancy, peak concentration for the centrally positioned nucleus is always less than the average randomly positioned nucleus and the randomly distributed structure. In addition, up to this point, the reaction trajectories for the randomly positioned nuclei and the randomly distributed internal clutter are not statistically significantly different from each other. At 50 % we see a discontinuity in behaviour, with the difference between randomly positioned nuclei and randomly distributed structure becoming statistically significant. At this SC ratio we also see the centrally positioned nucleus achieve a higher peak concentration than the randomly positioned structure.

CHAPTER 7

Quantifying the responses of simulated biological pathways to morphological changes of the cell

In this chapter we model how processes that change the shape of the cell alter cellular dynamics. We divide the simulations into two motifs, those that model a process in which volume is conserved, filamentation, and those in which volume is not conserved, hypertrophy and atrophy. First, a model *coccus*, *bacillus*, *spirillum* a non-standard cell shape have their dynamics simulated at different stages of hypertophy and atrophy. Second a model *bacillus* is allowed to undergo fillamentation with reaction dynamics recorded at several stages.

7.1 Hypertrophy and Atrophy

We begin our investigation of morphological change by examining hypertrophy and atrophy in a virtual *coccus*, *bacillus*, *spirillum* and abstract fried-egg shaped cell. This subsection details the different stages of these virtual cells and their sizes. Unaltered cell sizes and geometry are seen in Table 7.1.

For each stage of each cell, both the LVD and the BMR systems are simulated. The reaction systems are described in Table 7.2. In all simulations, the constituents of the virtual pathway are placed at a point directly next to the

Table 7.1: Geometry of the cell shapes used in these experiments

Cell	Geometry
Coccus	Diameter: $50\mu m$
Bacillus	Diameter: $25\mu m$ Length: $150\mu m$
Spirillum	Diameter: $28\mu m$ Length: $168\mu m$ Winding diameter: $16\mu m$

Table 7.2: Description of the systems used in these simulations. Neumann boundary conditions. $\Delta t = 10^{-14}s$, $\Delta x = 10^{-6}m$, $\alpha = 1.5 * 10^{12}$, $\beta = 1 * 10^{12}$, $\gamma = 3 * 10^{12}$, $k = 1.4^{12}s^{-1}$

System	Reaction Equation	Initial Conditions	Diffusion constant of all chemicals (mum^2/s)
BMR	$\frac{dA}{dt} = -k * A * B$ $\frac{dB}{dt} = -k * A * B$ $\frac{dC}{dt} = k * A * B$	$A = 10$ $B = 10$ $C = 0$	2.0
LVD	$\frac{dpred}{dt} = pred * (\alpha - \beta * prey)$ $\frac{dprey}{dt} = -prey(\gamma - \delta * pred)$	$pred = 10$ $prey = 10$	0.05

left hand inner membrane of the model cell. We call this location the "origin point" or OP. Exactly opposite this point, on the right hand side of the cell, we measure the system by recording time series of constituent concentrations. This location is known as the measurement point or MP. The locations of these two points within each stage of each cell are represented by a red point (the OP) and a green point (the MP) in the selected examples of virtual cell stages that are shown in this chapter.

Results in this chapter take two forms: graphs of concentration timeseries and concentration gradients. Time series show concentration information at one point (MP) in the cell over a specific time interval. Conversely, heatmaps show concentration at many points in the cell but only at a single point in time. The heatmaps in this work show concentration information in a 2D slice through the virtual cell. Concentration gradients are captured in the same plane as the source and measurement sites. Heatmaps are used to represent these concentration gradients.

When investigating hypertrophy and atrophy in this section, simulation results are presented in this order: First we examine timeseries measured at MP at the different stages of cell size, then (for selected simulations) we show the concentration gradients measured as the chemical of interest arrives at MP. These results are collected in the virtual *coccus*, then the *bacillus*, then the *spirillum* and finally the fried-egg shaped cell. This entire processes is itself repeated once for the LVD system and once for the BMR system.

Coccus stage	Volume
Stage 1	v_c
Stage 2	$1.25v_c$
Stage 3	$1.5v_c$
Stage 4	$1.75v_c$
Stage 5	$2.0v_c$ (dividing)
Stage 6	$2.0v_c$
Stage 7	$0.75v_c$
Stage 8	$0.5v_c$

Table 7.3: Sizes of the stages of the virtual *coccus*

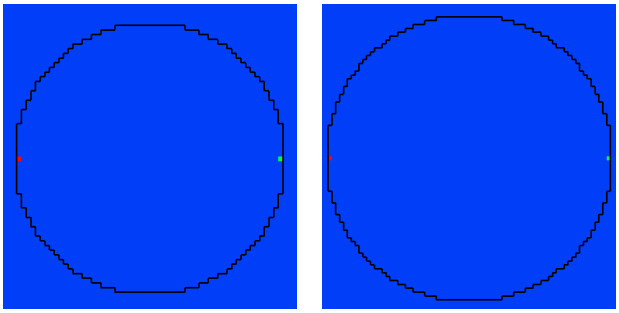
Bacillus stage	Volume
Stage 1	v_b
Stage 2	$2v_b$
Stage 3	$3v_b$
Stage 4	$4v_b$
Stage 5	$5v_b$
Stage 6	$6v_b$
Stage 7	$7v_b$
Stage 8	$8v_b$
Stage 9	$9v_b$
Stage 10	$10v_b$

Table 7.4: Sizes of the stages of the virtual *bacillus*

The virtual *coccus* is permitted to increase its volume, V_c, to twice its original size through four measurement points at $\{1.0, 1.25, 1.5, 1.75, 2.0\} \times V_c$, and decrease its volume to half its original size through two measurement points at $\{0.75, 0.5\} \times V_c$. Finally we model a dividing *coccus* of total volume $2 \times V_c$. The 8 stages of this virtual *coccus* are detailed in Table 7.3 and images of selected stages are seen in Figure 7.1.

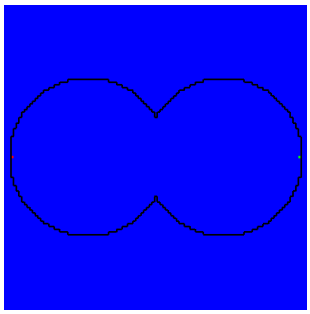
In simulations of the *bacillus*, the volume of the virtual cell is permitted to increase to 10 times the original size by increasing in length. Measurements of the cell are made at $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \times V_b$. This virtual cell and its 10 stages of hypertrophy are detailed in Table 7.4 and selected images of these stages are seen in Figure 7.2.

The virtual *spirillum* (of volume v_s) is allowed to increase to 4 times its original size by increasing in length. The cell is measured at 2 stages of hypertrophy: $\{2, 4\} \times v_s$. The stages of the *spirillum*'s hypertrophy are detailed in



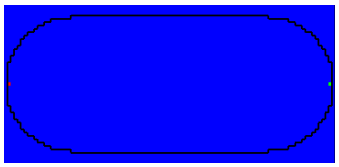
(a) Stage 1: virtual *Coccus* of volume v_c. (Radius of $29\mu m$)

(b) Stage 5: virtual *Coccus* of volume $2.0v_c$



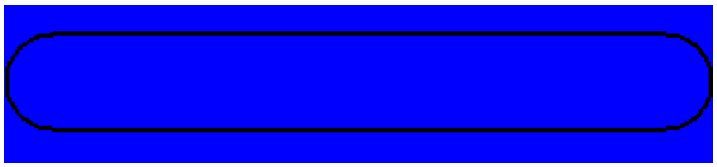
(c) Stage 6: dividing virtual *Coccus* of total volume $2.0v_c$

Figure 7.1: Initial configurations of selected stages of the virtual *coccus*



(a) Stage 1: virtual *bacillus* of volume v_b.

(b) Stage 2: virtual *bacillus* of length $2v_b$



(c) Stage 3: virtual *bacillus* of length $3v_b$

Figure 7.2: Initial configurations of selected stages of the virtual *bacillus*

Spirillum stage	Volume
Stage 1	v_s
Stage 2	$2v_s$
Stage 3	$4v_s$

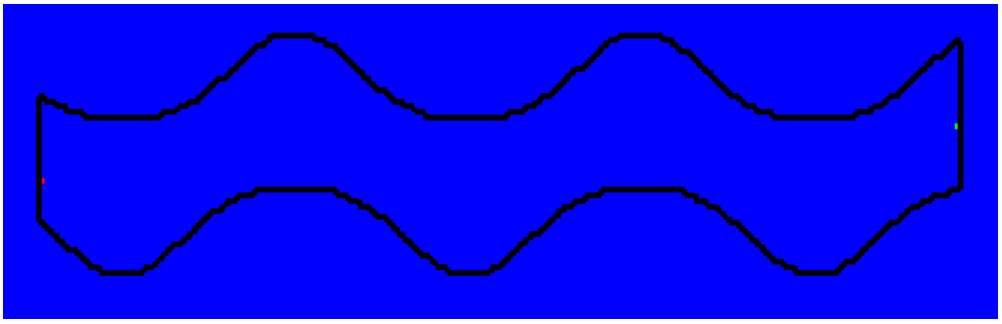
Table 7.5: Sizes of the stages of the virtual *spirillum*

Table 7.5 and selected images of these stages are seen in Figure 7.3.

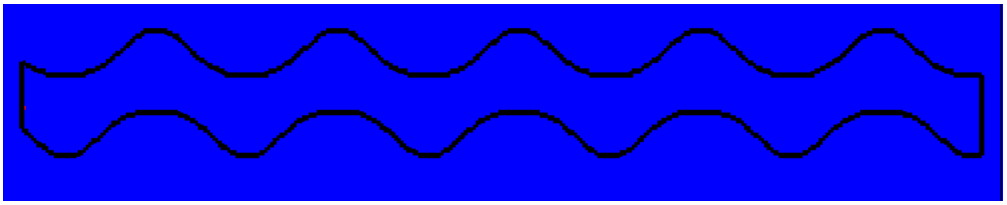
The virtual fried-egg shaped cell (of volume V_i) is permitted to increase to twice its original size, through measurements points at $\{1, 2\} \times V_s$. The cell is increased proportionally in all dimensions. This virtual cell and its 2 stages of hypertrophy are detailed in Table 7.6 and images of these stages are seen in Figure 7.4.

Fried-egg cell stage	Volume
Stage 1	V_s
Stage 2	$2V_s$

Table 7.6: Sizes of the stages of the virtual fried-egg cell

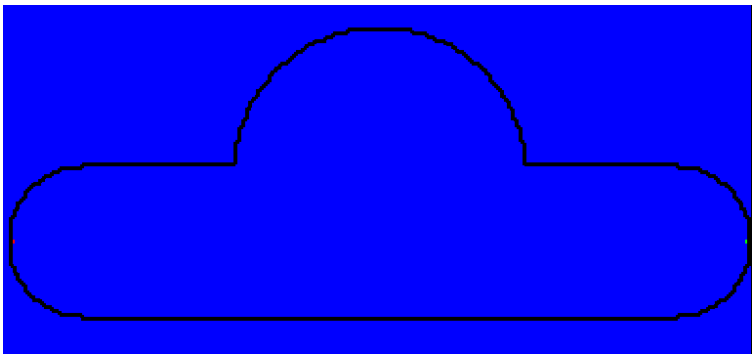


(a) Stage 1: virtual *spirillum* of volume v_s.

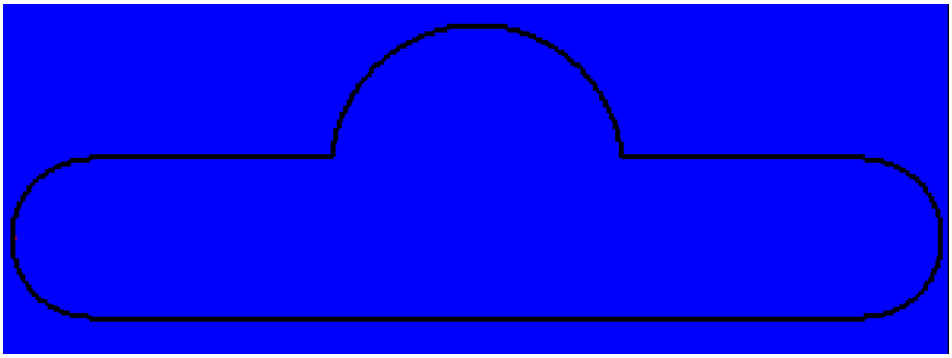


(b) Stage 2: virtual *spirillum* of volume $2v_s$

Figure 7.3: Initial configurations of selected stages of the virtual *spirillum*



(a) Stage 1: virtual fried-egg cell of volume v_i.



(b) Stage 2: virtual fried-egg cell of volume $2v_i$

Figure 7.4: Initial configurations of a virtual irregularly shaped cell in different stages of hypertrophy.

7.1.1 The LVD system

We begin by simulating the Lotka-Volterra-Diffusion (LVD) system within the virtual *coccus*. Figure 7.5 shows the timeseries of *predator* concentration at all stages of the virtual *coccus*'s hypertrophy and atrophy (detailed in Table 7.3).

In Figure 7.6 we show time series from the smallest and largest stages of the virtual *coccus*, shifted so that they begin at the same time making concentration behaviour easier to observe.

Next we compare the behaviour of the virtual *coccus* just before division (Figure 7.1(b)) and just before complete separation (Figure 7.1(c)). The results of this simulation are seen in Figure 7.7.

Figure 7.8 compares the concentration gradients formed by the different stages of hypertrophy within the virtual *coccus*.

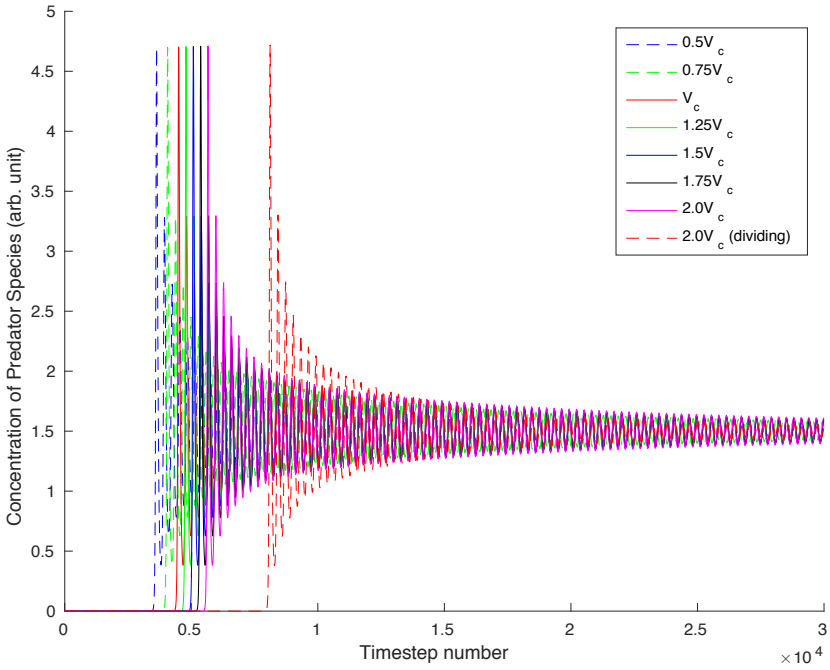


Figure 7.5: Time series of *predator* species at all stages of the virtual *coccus*. It is immediately obvious that as distance between OP and MP increases the time of arrival increases. Differences in concentration behaviour are harder to discern.

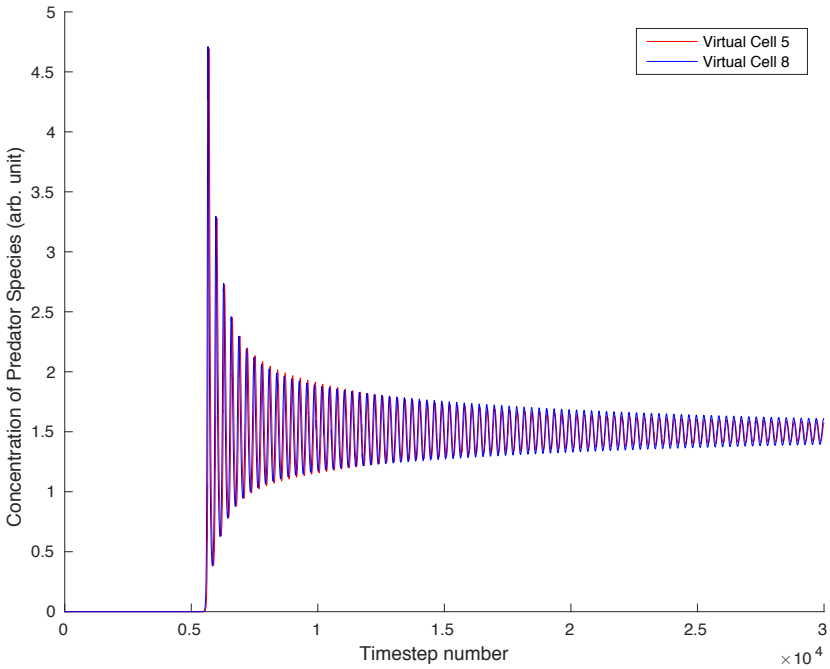


Figure 7.6: Time series of *predator* species at stages 5 ($2v_c$, dividing) and 8 ($0.5v_c$) (Table 7.3) of the virtual *coccus*. This figure contrasts concentration trajectory behaviour in the largest and smallest stages of the virtual *Coccus*. Time series are shifted to have identical arrival times so that concentration behaviour is easier to observe. Once shifted, it is clear that there is little difference in concentration behaviour between the smallest and the largest stages of the virtual *coccus*, aside from a slight difference in amplitude.

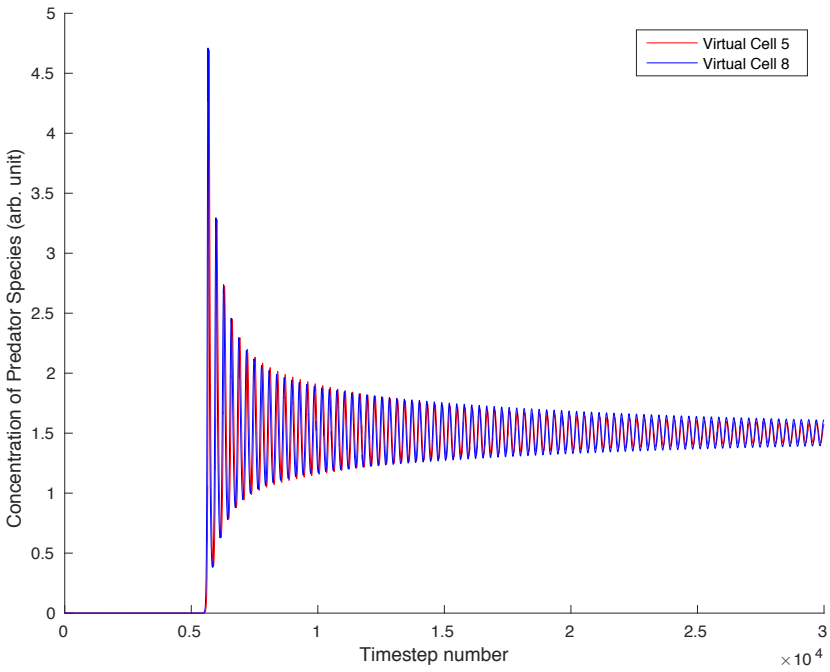
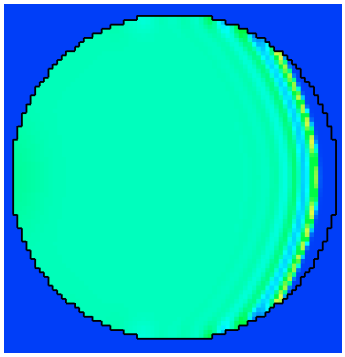
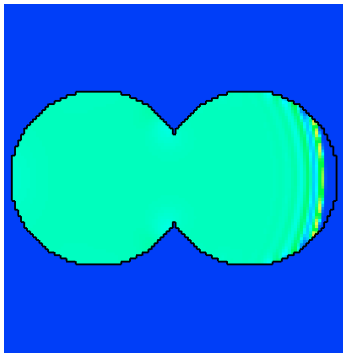


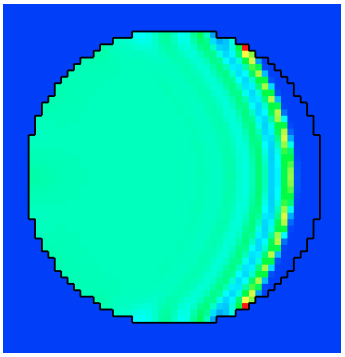
Figure 7.7: Time series of *predator* species at stages 5 ($2v_c$, dividing) and 6 ($2v_c$) (Figure 7.1) of the virtual *coccus*. This figure contrasts concentration trajectory behaviour in a pre-division and dividing virtual cell. Time series are shifted to have identical arrival times so that concentration behaviour is easier to observe.



(a) Concentration gradient of *predator* species within a stage 5 ($2v_c$, dividing) virtual *coccus*



(b) Concentration gradient of *predator* species within a stage 6 ($2v_c$) virtual *coccus*



(c) Concentration gradient of *predator* species within a stage 8 ($0.5v_c$) virtual *coccus*

Figure 7.8: Concentration gradients of *predator* species within selected stages of a virtual *coccus*. In all three stages the shape of wavefront of concentration remains similar. The smallest stage (8) displays a higher concentration at the edges of this travelling wave of concentration.

Having examined the behaviour of the LVD system inside an enlarging and dividing *coccus* we turn our attention to an enlarging *bacillus*. Figure 7.9 shows the time series of *predator* concentration at all stages of the virtual *bacillus* (detailed in Table 7.4).

Next we employ the same technique of overlaying results as in Figure 7.6 to see if there are any differences in concentration behaviour. Figure 7.10 shows this comparison.

Next, in Figure 7.11 we compare the concentration gradients formed at the different stages of an enlarging *bacillus*.

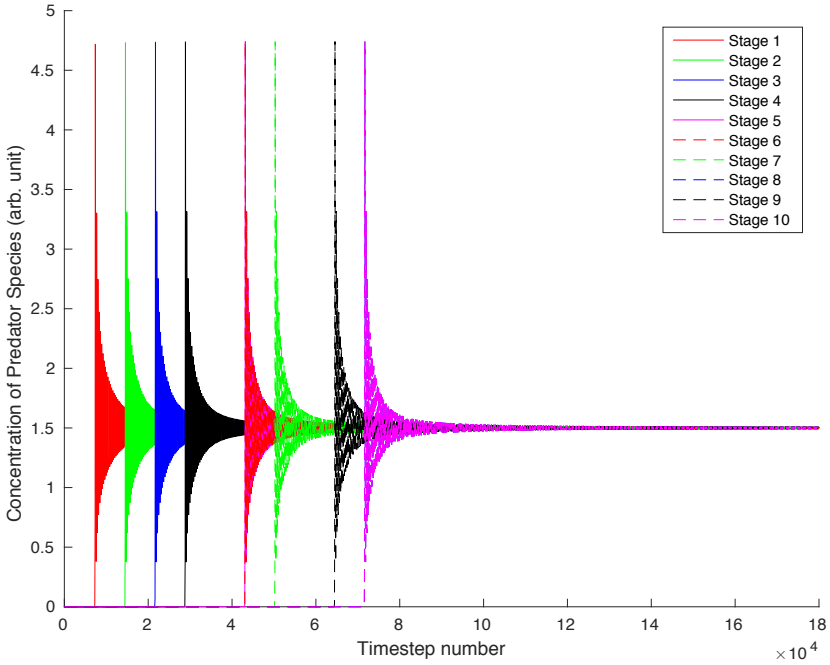


Figure 7.9: Time series of *predator* species at all stages of the virtual *bacillus* as seen in Table 7.4. As the length of the *bacillus* increases, the time of arrival of the *predator* species also increases.

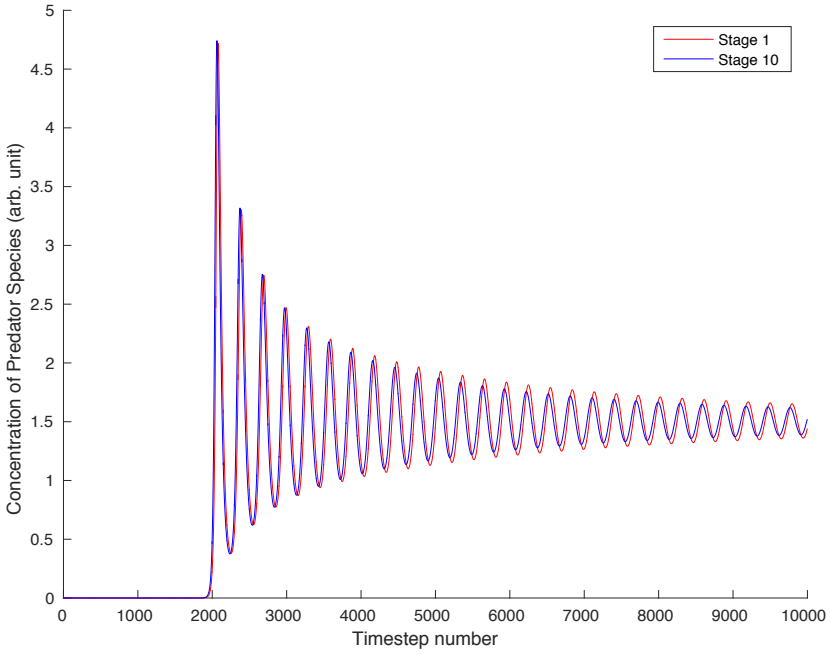
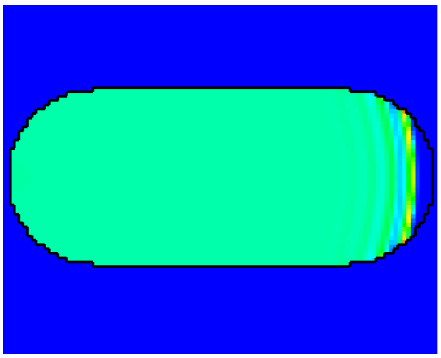
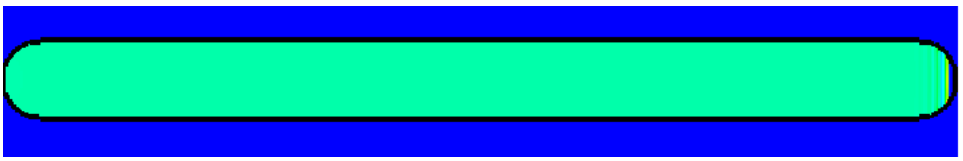


Figure 7.10: Time series of *predator* species at stages 1 (v_b) and 10 ($10v_b$) of the virtual *bacillus*. Time series are shifted to have identical arrival times so that concentration behaviour is easier to observe. Differences between these two timeseries are small, and so further work may be required to identify if the effect of lengthening is significant.



(a) Concentration gradient of *predator* species within a stage 1 (v_b) virtual *bacillus*



(b) Concentration gradient of *predator* species within a stage 5 ($5v_b$) virtual *bacillus*

Figure 7.11: Concentration gradients of *predator* species within selected stages of a virtual *bacillus*. Nearly identical concentration gradients and wavefront shapes are observed in both virtual cells.

The *spirillum* is the next cell shape explored. Once again we begin by recording time series of *predator* concentration at the different stages of the *spirillum* (Table 7.5). It is clear that as the length of the cell increases the time of arrival of *predator* species also increases (Figure 7.12).

As in the previous cell shapes, the time series from all stages of the *spirillum* are then overlayed such that they all have the same arrival time; this allows concentration behaviour to be compared with greater ease. This analysis can be seen in Figure 7.13.

As with the previous cell shapes, the concentration gradients are measured. These results are presented in Figure 7.14. Whilst the values of concentration are similar at both stages, small differences in cell shape result in a differently shaped wavefront of *predator* species.

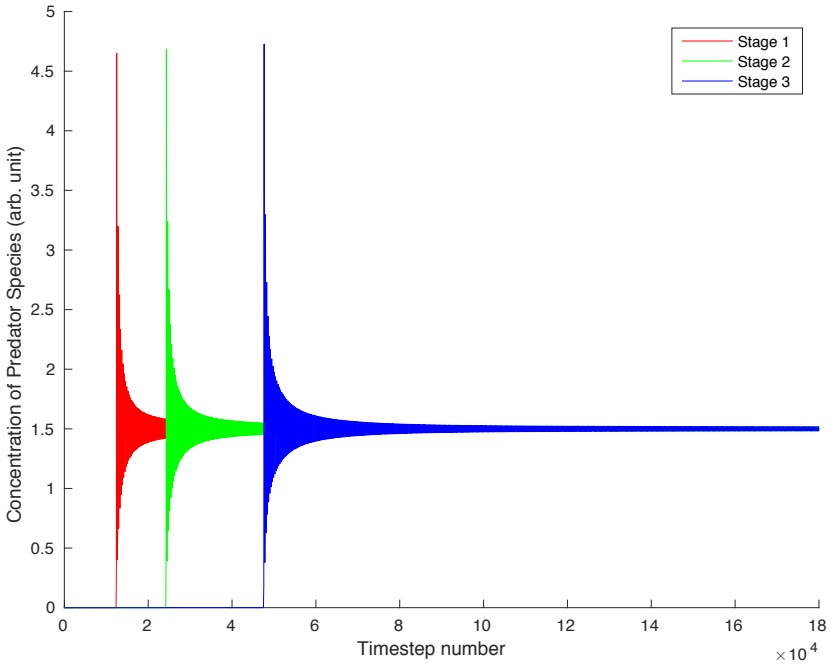


Figure 7.12: Time series of *predator* species at all stages of the virtual *spirillum*

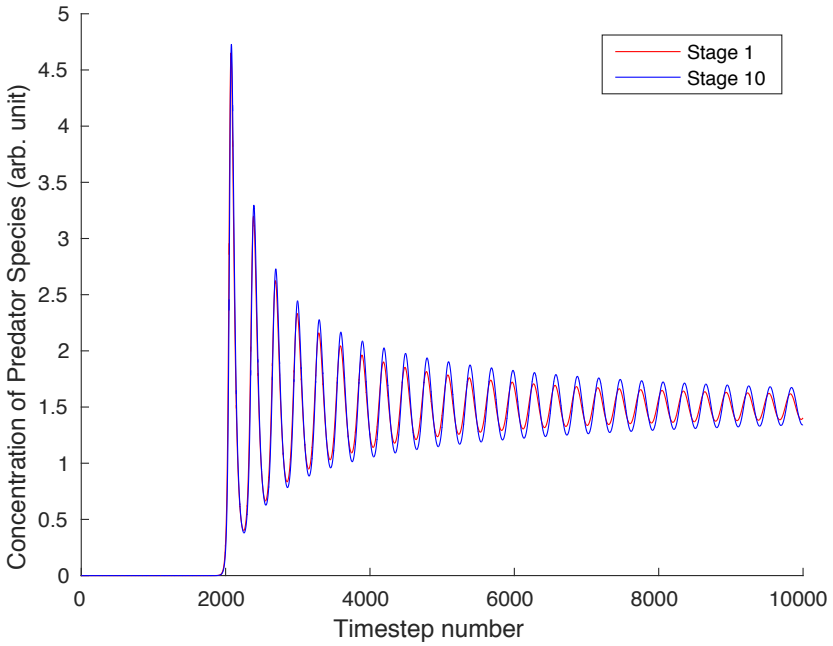
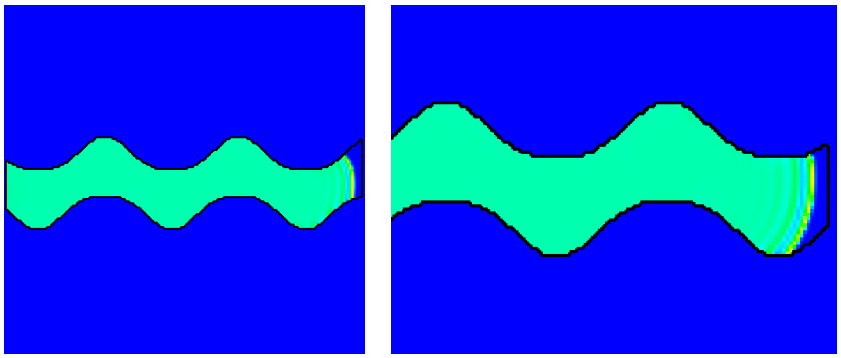


Figure 7.13: Time series of *predator* species in the largest ($4v_s$) and smallest (v_s) stages of the virtual *spirillum*. Time series are shifted to have identical arrival times so that concentration behaviour is easier to observe. The smaller *spirillum* results in a concentration profile with a slightly smaller amplitude. It could be argued that the differences in concentration are negligible. Further experimentation is needed.



(a) Concentration gradient of *predator* species within a stage 1 (v_s) virtual *spirillum*

(b) Concentration gradient of *predator* species within a stage 3 ($4v_s$) virtual *spirillum*

Figure 7.14: Concentration gradients of *predator* species at selected stages of the virtual *spirillum*

Finally we turn our attention to hypertrophy in the fried-egg cell depicted in Figure 7.4. As in previous simulations, we begin by plotting the time series as recorded at MP in Figure 7.15.

We repeat the analyses from previous cell shapes by aligning the arrival times from both time series in Figure 7.16.

Concentration gradients in both sizes of fried-egg cell can be seen in Figure 7.17.

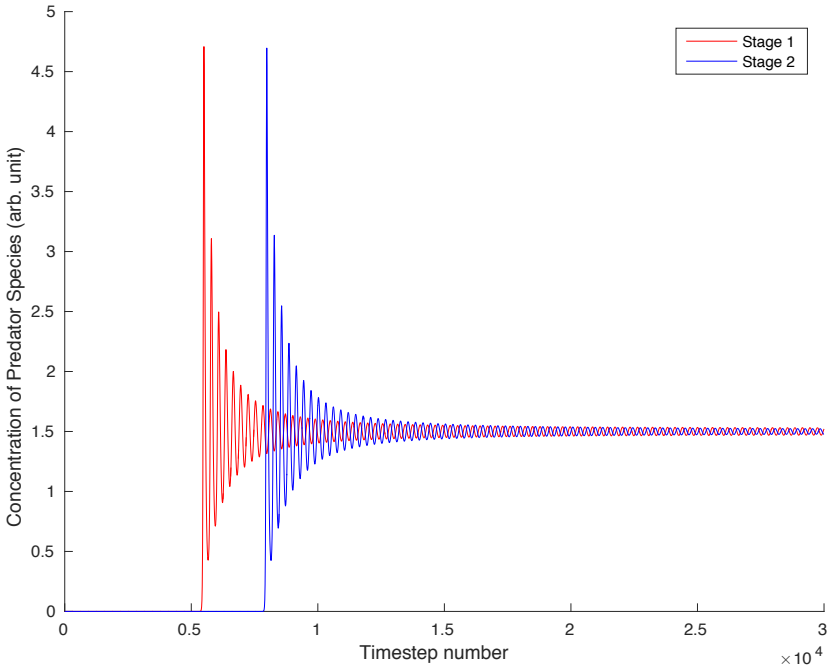


Figure 7.15: Time series of *predator* species at all stages of the virtual fried-egg cell as seen in Figure 7.4. The larger virtual cell results in later arrival time.

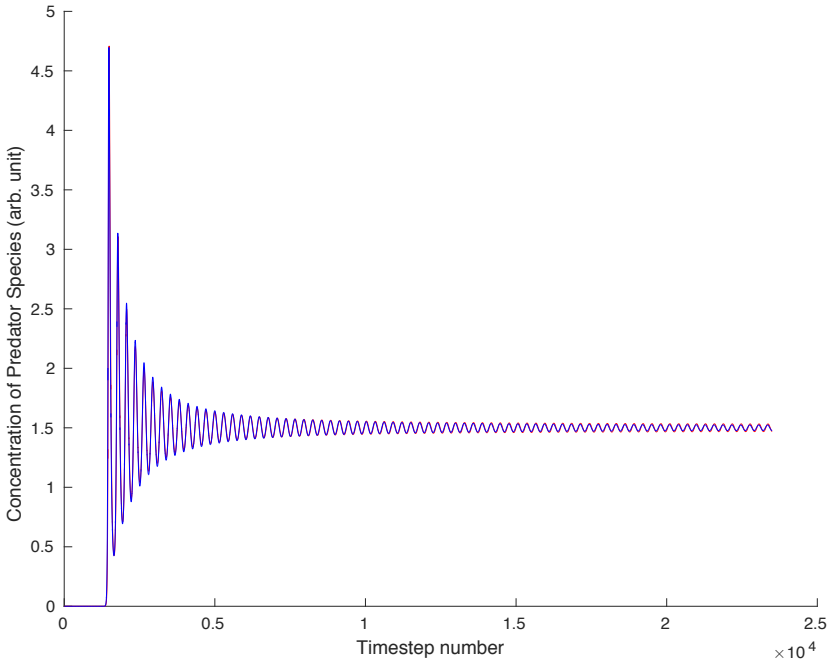
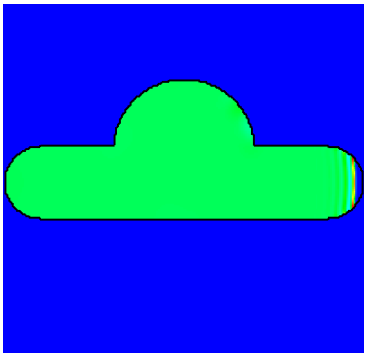
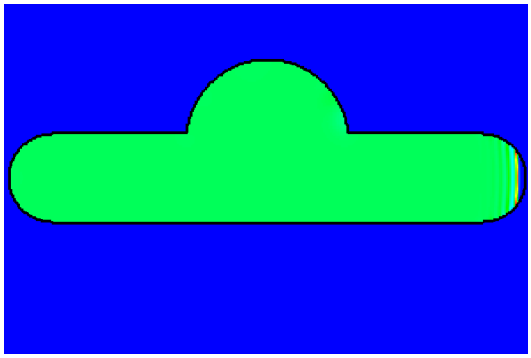


Figure 7.16: Time series of *predator* species in the largest ($2v_i$) and smallest (v_i) stages of the virtual fried egg cell as seen in Figure 7.4. Time series are shifted to have identical arrival times so that concentration behaviour is easier to observe. Differences in concentration are slight.



(a) Concentration gradient of *predator* species within a stage 1 (v_i) virtual fried-egg cell



(b) Concentration gradient of *predator* species within a stage 2 ($2v_i$) virtual fried-egg cell

Figure 7.17: Concentration gradients of *predator* species within the stages of the virtual fried-egg cell. Both stages of the cell result in similar concentration gradients and wavefront shapes.

7.1.2 The BMR system

We repeat all of the simulations in the Subsection 7.1.1 but instead we substitute the LVD system for the Bi-Molecular Reaction (BMR) system.

The virtual *coccus* is the first cell explored. The time series from all stages of the virtual *coccus* are displayed in Figure 7.18.

Next we compare the behaviour of the *coccus* just before division (Figure 7.1(b)) and just before complete separation (Figure 7.1(c)). The results of this simulation are seen in Figure 7.19.

Concentration gradient heatmaps are also shown in Figure 7.20.

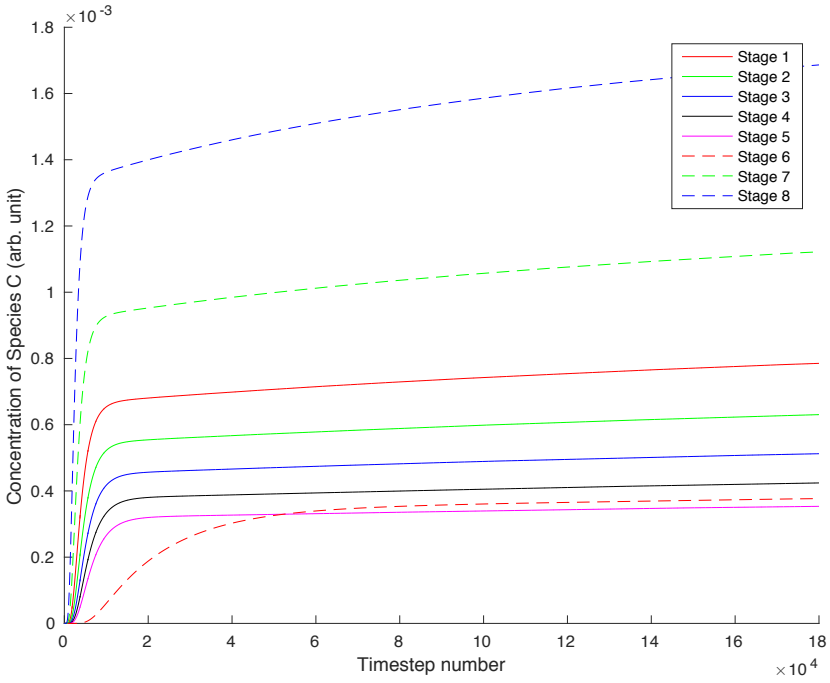


Figure 7.18: Time series of species C in all stages of the virtual *coccus* as seen in Figure 7.1. As the distance between OP and MP increases, the time of arrival of species C at MP also increases. As expected, smaller cells result in higher concentrations, which is contrast to results from the LVD system. This is due to the inability of any of the reactants in the BMR system to make more of themselves, unlike the LVD system.

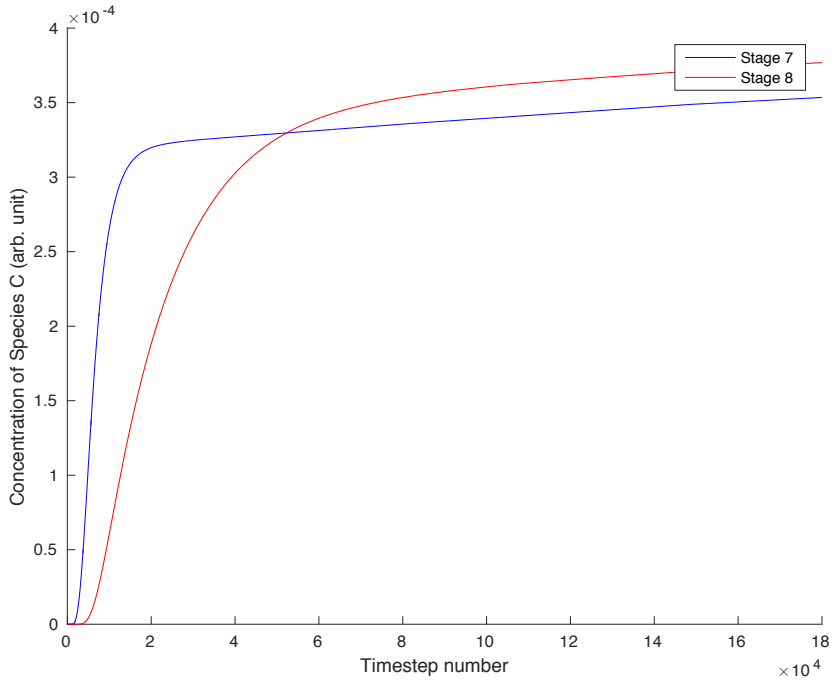
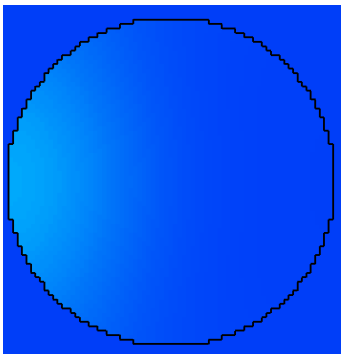
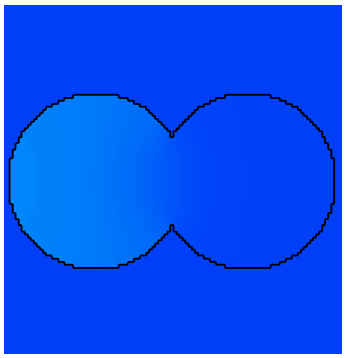


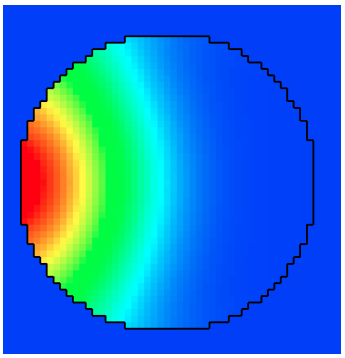
Figure 7.19: Time series of species C in a dividing cell (stage 5 $2v_c$, dividing) and a pre-division virtual cell (stage 6, $2v_c$). A clear difference in concentration trajectory between the two cells is seen, despite having identical volumes.



(a) Concentration gradient of species C within a stage 5 ($2v_c$, dividing) virtual *coccus*



(b) Concentration gradient of species C within a stage 6 ($2v_c$) virtual *coccus*



(c) Concentration gradient of species C within a stage 8 ($0.5v_c$) virtual *coccus*

Figure 7.20: Concentration gradients of species C within selected stages of a virtual *coccus*. Concentration gradient differs between all stages, although the shape of the wavefront remains the same.

Next we turn our attention to an enlarging *bacillus*. Figure 7.21 shows the time series of species C concentration at all stages of the virtual *bacillus* (detailed in Table 7.4).

Figure 7.22 shows concentration gradients captured in stages 1 and 5 of the virtual *bacillus*. Predictably in the larger virtual cell the concentration is lower. In both images the shape of the wavefront is identical.

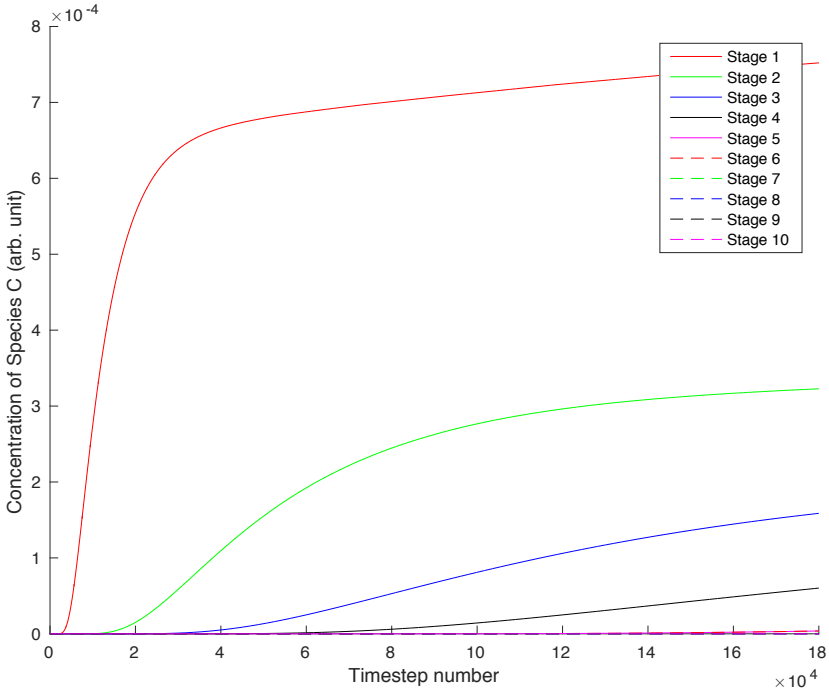
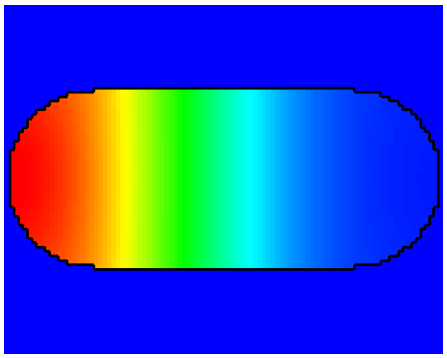


Figure 7.21: Time series of species C in all stages of the virtual *bacillus*. As the length of the *bacillus* increases, the time of arrival of species C also increases. Peak concentrations and reaction trajectories are markedly different between the different stages of the *bacillus*. As the cell becomes longer the peak concentration drops and the reaction trajectories become "flattened". As in Figure 7.18 the differences in reaction trajectory are much more apparent than in the LVD system. Smaller stages of the virtual cell result in higher peak concentrations and faster progression of reactions.



(a) Concentration gradient of species C within a stage 1 (v_b) virtual *bacillus*

(b) Concentration gradient of species C within a stage 5 ($5v_b$) virtual *bacillus*

Figure 7.22: Concentration gradients of species C within selected stages of a virtual *bacillus*

We continue our analysis of the BMR system by examining its behaviour within a *spirillum* of changing size. The stages of the virtual *spirillum* are detailed in Table 7.5. Figure 7.23 shows the time series of species C concentration in all stages of the virtual *spirillum*.

Figure 7.24 shows concentration gradients in different sizes of *spirillum*.

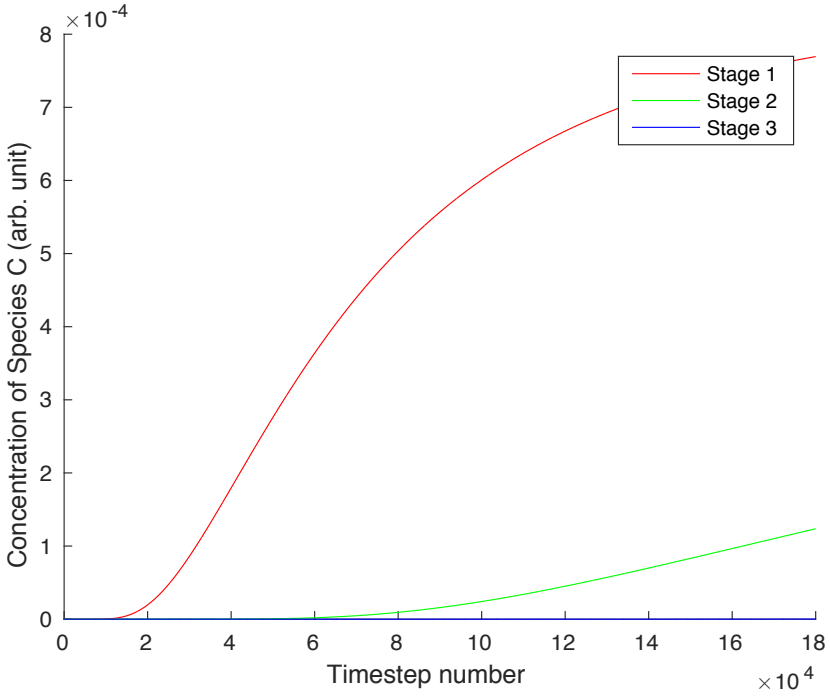
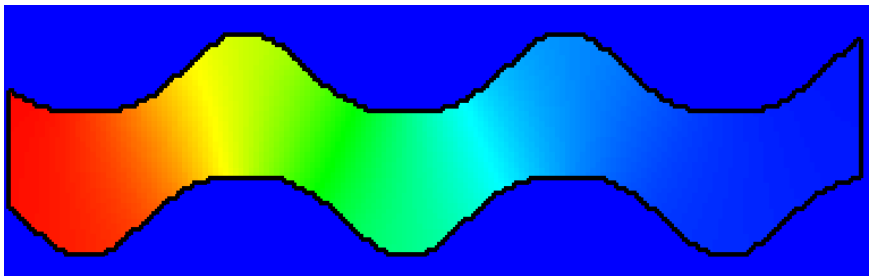
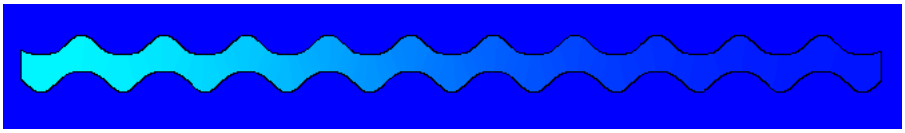


Figure 7.23: Time series of species C in all stages of the virtual *spirillum* as seen in Figure 7.3. As the length of the *spirillum* increases, the time of arrival of species C species also increases. Differences in concentration trajectory are also apparent, with stages of smaller volume giving rise to faster reactions and higher peak concentrations.



(a) Concentration gradient of species C within a stage 1 (v_s) virtual *spirillum*



(b) Concentration gradient of species C within a stage 3 ($4v_s$ virtual *spirillum*

Figure 7.24: Concentration gradients of species C within selected stages of a virtual *spirillum*. Differently sized virtual cells result in different concentration gradients.

This process is repeated one final time for the fried-egg shaped cell depicted in Table 7.6. We begin by plotting the time series of species C in Figure 7.25. The concentration gradients for this cell shape are shown in Figure 7.26.

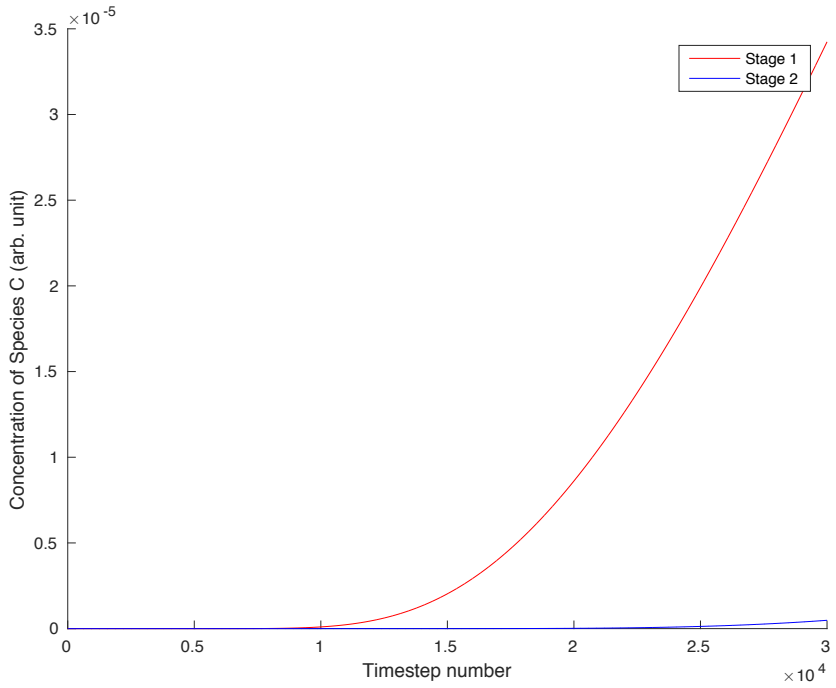
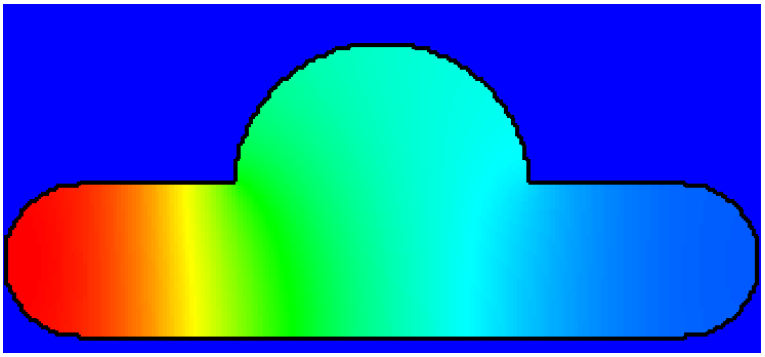
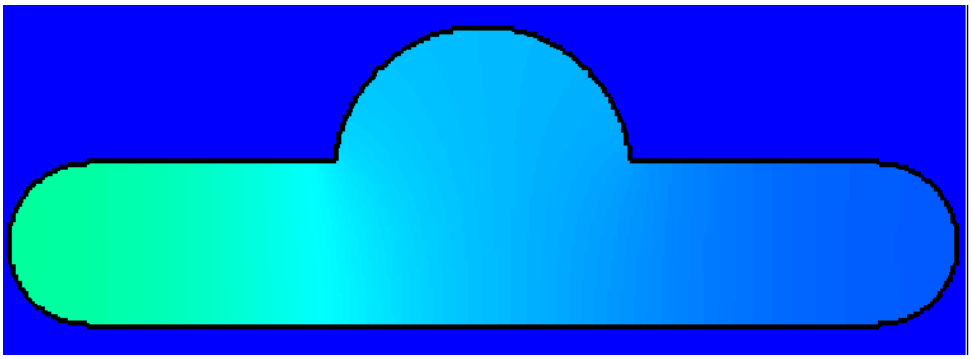


Figure 7.25: Time series of species C in all stages of the virtual fried-egg cell. The pattern of increasing volume within the BMR system repeats itself with larger cells showing later arrival time, a slower reaction rate and a lower peak concentration.



(a) Concentration gradient of species C within a stage 1 (v_i) virtual fried-egg cell



(b) Concentration gradient of species C within a stage 2 ($2v_i$) virtual fried-egg cell

Figure 7.26: Concentration gradients of species C within selected stages of a virtual fried-egg cell. The smaller cell results in greater concentrations, but in both stages the wavefront shape is the same.

7.2 Elongation

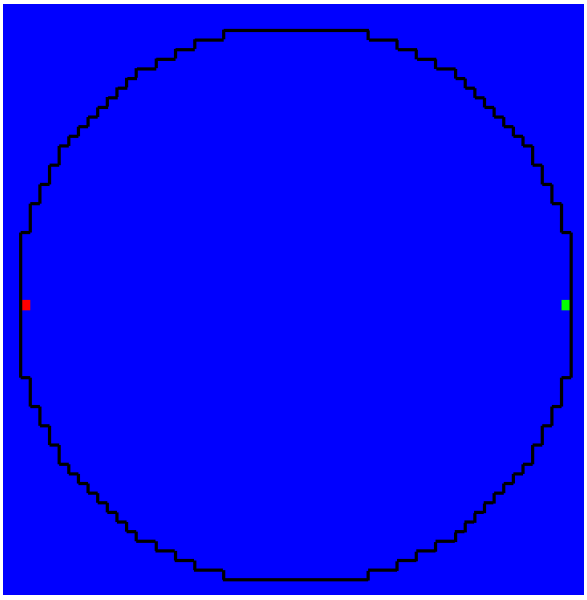
The final set of simulations in this chapter examines reaction system behaviour within a virtual *bacillus* undergoing elongation in response to an environmental stimulus such as nutrient deficiency. To quantify the degree of elongation, we define the "'Fatness" or "f" parameter which is equal to the new height of the elongated *bacillus* divided by its old height (Equation 7.1).

$$f = \frac{height_{new}}{height_{old}} \tag{7.1}$$

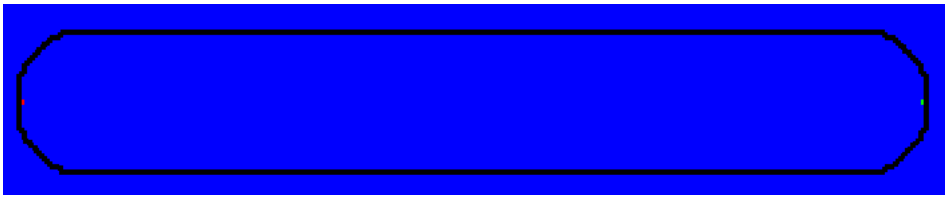
simultaneously:

$$volume_{old} = volume_{new} \tag{7.2}$$

No matter the height or length of the bacteria, the volume is kept constant. elongation is a continuous, rather than a discrete process, however here we measure reaction trajectories at 4 stages as show in Figure 7.27. These four stages represent "fatness" constants $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$. Time series at all of the stages in Figure 7.27 are measured when simulating both the LVD and BMR reaction systems. As in the previous examples, time series are measured at MP and concentration gradients are captured when a specific chemical species reaches MP. Concentration gradient images are 2D slices of the 3D cell. In all gradients, the red subvolume denotes the OP and green subvolume the MP.

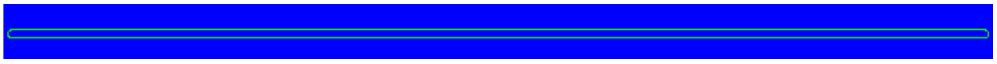


(a) Stage 1: virtual cell of f ratio 1



(b) Stage 2: virtual cell of f ratio 1/2

(c) Stage 3: virtual cell of f ratio 1/4



(d) Stage 4: virtual cell of f ratio 1/8

Figure 7.27: Figure shows initial configurations of a virtual *bacillus* in different stages of elongation. Images are a 2D slice of the 3D cell. In all images the red subvolume denotes the OP and green subvolume the MP

7.2.1 LVD

First, we simulate the LVD system within the four stages of the elongating *bacillus*. The results from this set of simulations are shown in Figure 7.28.

It is unclear whether concentration trajectory is altered by different f ratios in Figure 7.28. Therefore, we re-plot the results from Figure 7.28 shifting the timeseries so that they posses the same starting point - this analysis is seen in Figure 7.29.

Finally, we compare concentration gradients at all stages of elongation. These results are seen in Figure 7.30 and the peak concentrations at all points in the virtual cell are shown in Table 7.7.

Table 7.7: Peak concentration of *predator* species in the elongating virtual cell

f ratio	Peak concentration, *predator* species
1	4.69
$\frac{1}{2}$	3.928
$\frac{1}{4}$	4.361
$\frac{1}{8}$	4.44

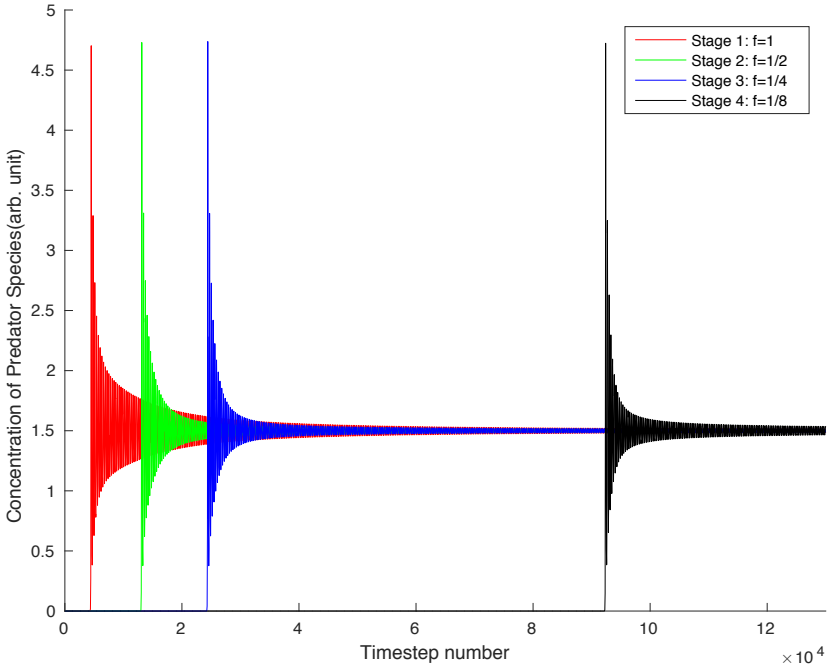


Figure 7.28: Figure showing the reaction trajectories at all stages of the virtual cell as seen in Figure 7.27. Despite conserving the cell's total volume, the longer the cell, the later the arrival time.

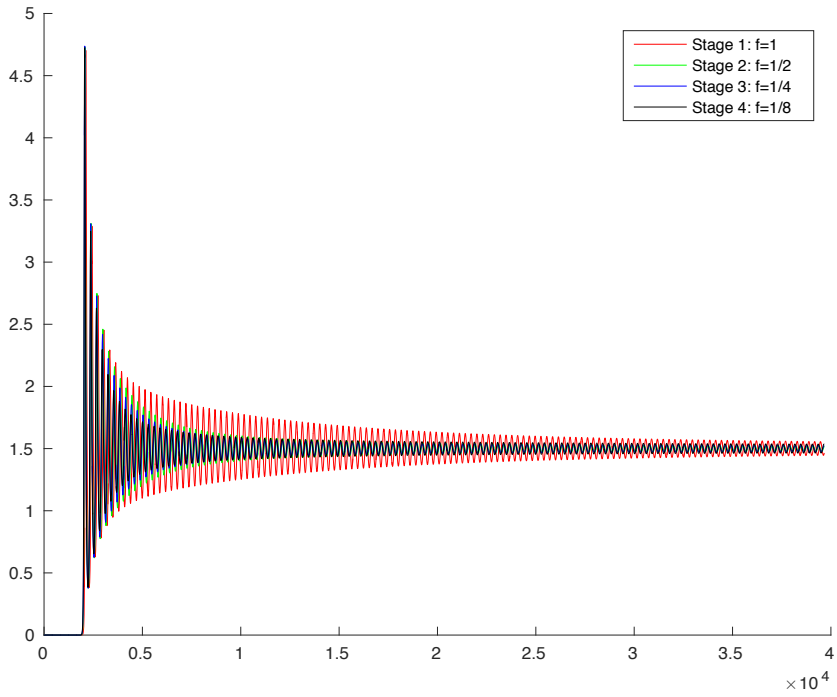
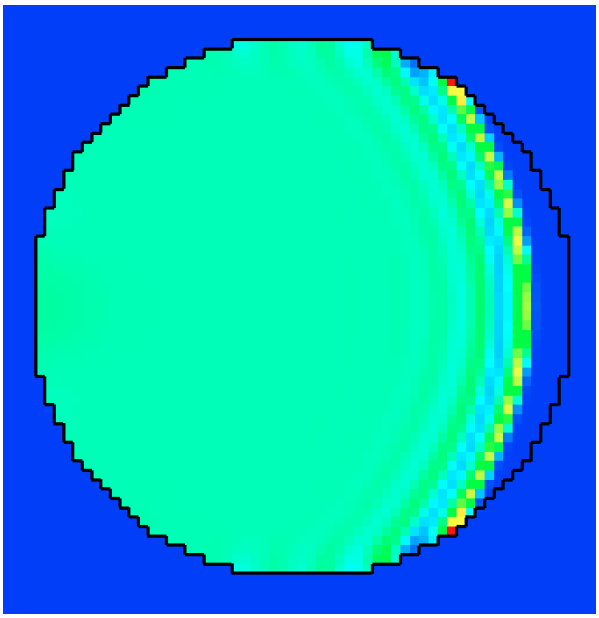
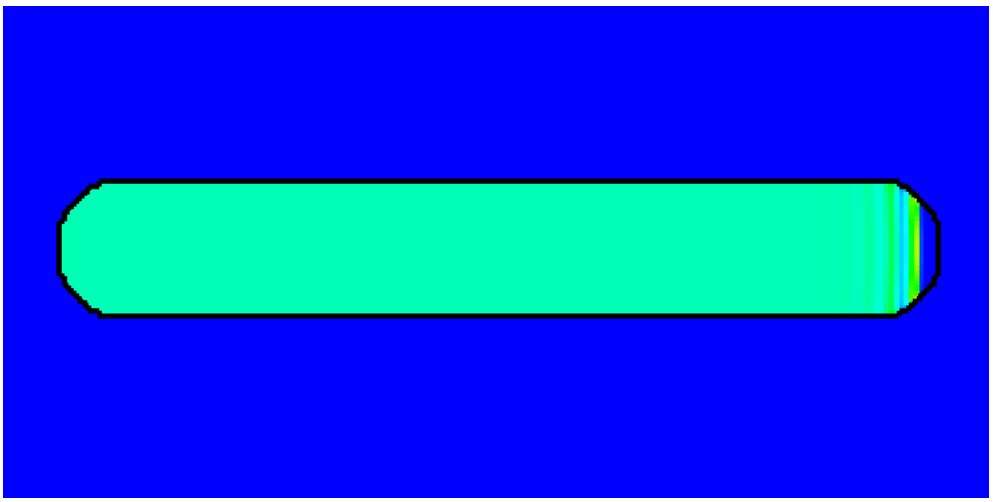


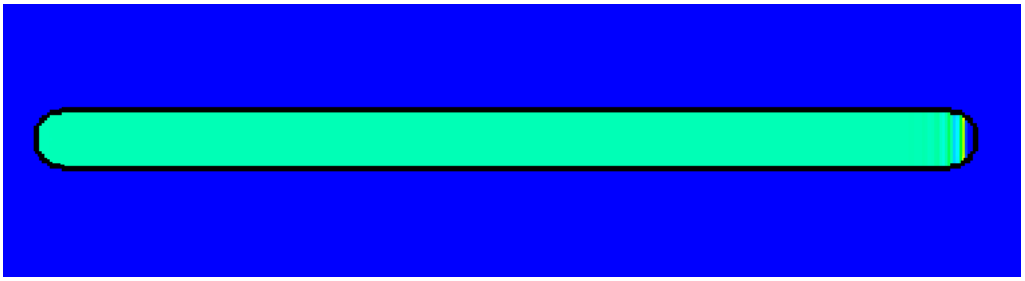
Figure 7.29: Figure showing the reaction trajectories at all stages of the moving virtual cell as seen in Figure 7.27. Timeseries shifted so that they have same arrival time. The smaller the f ratio, the more the *predator* species concentration time series is damped. As the f ratio decreases, the effect on damping behaviour decreases.



(a) Stage 1: virtual cell of f ratio 1



(b) Stage 2: virtual cell of f ratio 1/2



(c) Stage 3: virtual cell of f ratio 1/4

Figure 7.30: Figure shows concentration gradients in a virtual *Coccus* in different stages of elongation. The virtual cell at f ratio 1/8 is not shown as the width of the cell is too small to be meaningful. As the *coccus* elongates into a *bacillus*, the shape of the wavefront changes.

7.2.2 BMR

We now repeat the elongation experiments seen in Figure 7.28, substituting the LVD system for the BMR system. The results from this set of time series are seen in Figure 7.31.

Figure 7.32 shows concentration gradients at the different stages of elongation process as species C arrives at MP. Table 7.8 records the peak concentrations in the cell at this point.

Table 7.8: Peak concentration of species C in the elongating virtual cell. A non linear response of concentration to f ratio is observed.

f ratio	Peak concentration, species c
1	0.002322
$\frac{1}{2}$	0.001663
$\frac{1}{4}$	0.001848
$\frac{1}{8}$	0.002366

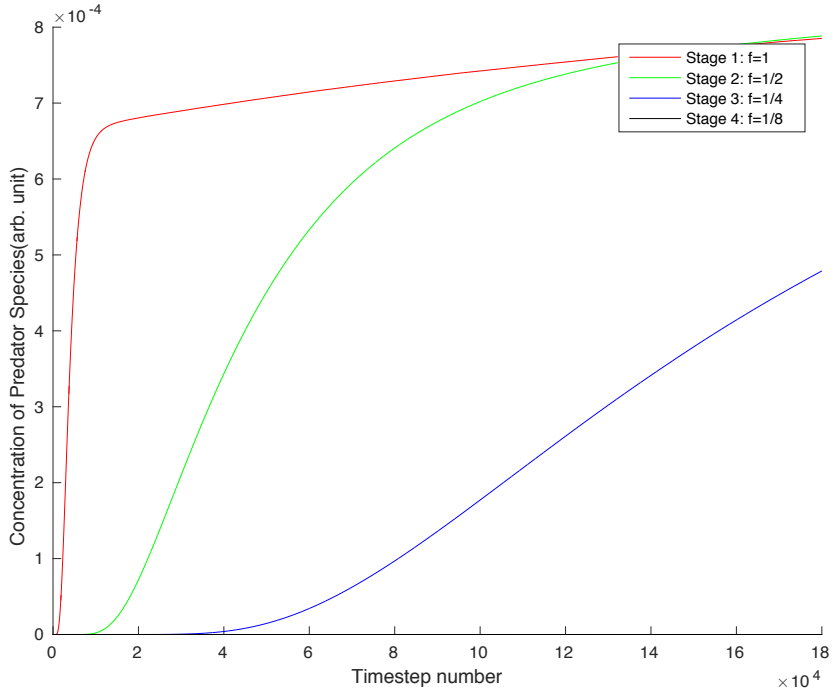
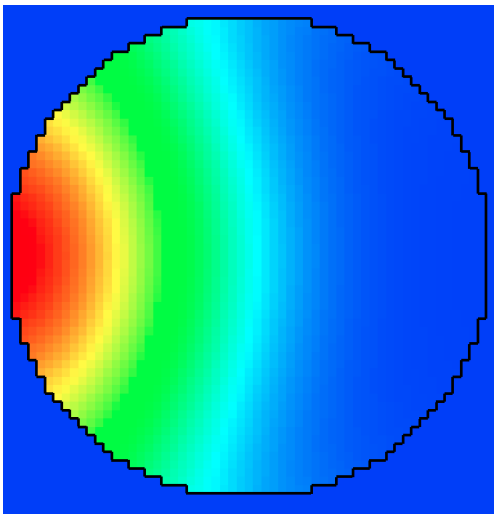
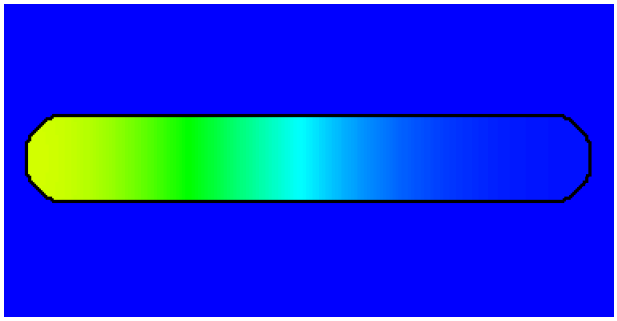


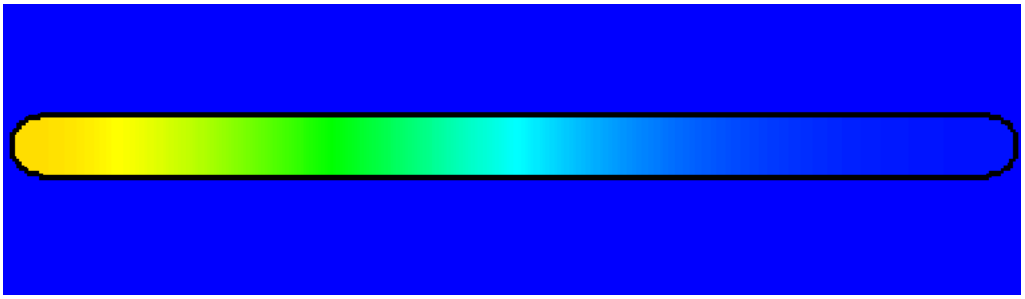
Figure 7.31: Time series of species C at all stages of the virtual cell as seen in Figure 7.27. Elongating the cell results in different reaction trajectories.



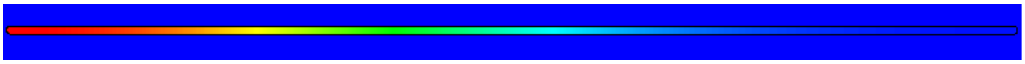
(a) Stage 1: virtual cell of f ratio 1



(b) Stage 2: virtual cell of f ratio 1/2



(c) Stage 3: virtual cell of f ratio 1/4



(d) Stage 4: virtual cell of f ratio 1/8

Figure 7.32: Figure shows concentration gradients in a virtual *Coccus* in different stages of elongation. Concentration gradients, peak concentration values and wavefront shape differ in these figures. A circular wavefront is presented in Figure 7.32(a) along with the highest peak concentration. Stages 2 (Figure 7.32(b)), 3 (Figure 7.32(c)) and 4 (Figure 7.32(d)) present straight wavefronts. As f ratio increases a non-linear response to peak concentration is observed (Table 7.8

7.3 Analysis of results

In this chapter, we have modelled how processes that change the shape of the cell alter cellular dynamics. We divided the simulations into two motifs, those that change shape and do not conserve volume and those that change shape and do conserve volume. When considering the effect of cell size upon reaction-diffusion systems, the choice of reaction system is very important.

Within the LVD system, very little difference in reaction trajectory or concentration gradient is achieved by increasing or decreasing the size of the virtual cell. When a change in cell size does cause the trajectory to change, there is a slight change in concentration timeseries amplitude. When considering processes that change the geometry of the cell, it would appear that only processes that change the *shape*, rather than the *size* of the virtual cell, have much of an impact on the LVD timeseries. The only major difference in concentration trajectory occurs when the *coccus* divides. This process changes the shape rather than just changing the volume. Peak concentrations are the only universally sensitive metric (as recorded in the peak concentration tables), varying in a non-intuitive fashion.

Behaviour in the BMR system is markedly different. This is due to the lack of feedback mechanism within the BMR system; neither species A, B or C can produce more of themselves without other reactants. As expected, smaller cells result in faster reactions and higher peak concentrations than larger cells. There are however, some unexpected exceptions to this; the dividing *coccus* results in a different reaction trajectory to the pre-division *coccus* despite being of identical volume. The lack of feedback behaviour has a similar effect on the concentration gradients. Larger cells result in lower concentrations in the heatmaps, but the wavefront shapes are similar. Reaction trajectory, wavefront shape and concentration gradient are all much more sensitive in the BMR system than the LVD system.

CHAPTER 8

Conclusions

In this thesis we have systematically explored the effect that spatial configuration has on simulated biological pathways. The three aspects of spatial configuration that have been explored are:

- external cell shape
- internal cellular organisation
- processes that change the shape of the cell

We have shown how the outcome of two abstract simulated pathways varies with each of these aspects of morphology. This chapter reviews the contributions made in this thesis with respect to the research questions in chapter 1, identifies conclusions that can be drawn across all three results chapters (5, 6, 7) and identifies directions for future work.

8.1 Contributions

Quantifying the responses of simulated biological pathways to common bacterial cell shapes

In chapter 4, we showed how the three common bacterial shapes *cocci*, *bacilli*, and *spirilli* alter the behaviour of two simulated biological pathways. Bacterial cells take on a bewildering variety of external shapes but the shapes listed above are the three most common. We showed which physical properties of the model pathway change when the cell shape changes; specifically, wavefront shape, peak concentration, concentration gradient and concentration distribution are all affected by the shape of the cell. Furthermore, we showed that these responses not

only depend on the shape but on the reaction system themselves. Concentration gradients, peak concentration and concentration distribution display markedly different patterns of response to the different pathways. However, the shape of the travelling wave of concentration (the wavefront shape) is comparable across cells of the same type when the choice of simulated pathway is different. Finally we demonstrate the resistance of a third type of simulated reaction pathway to discontinuities in the virtual cell environment. From these simulations we can draw five conclusions.

First, different bacterial shapes alter the behaviour of reaction-diffusion systems, even when controlling for internal volume. In both reaction systems, changing the shape of the cell resulted in changes in all of the recorded metrics.

Second, some reaction systems are more sensitive to external cell shape than others; in the Lotka-Volterra-Diffusion (LVD) system, changing the shape of the cell has a large impact on the spatial pattern of concentration, but has less impact on absolute concentration. On the other hand, the Bi-Molecular Reaction (BMR) pathway produces identical *patterns* of concentration with far greater difference in absolute concentrations. As a final example of this phenomenon, the Turing system differs very little in terms of pattern of concentration or indeed in terms of absolute concentrations.

Third, wavefront shape is, at least in part, dictated by the external shape of the cell (rather than choice of reaction system), as wavefront shapes were strikingly similar across the reaction systems in identical cell shapes.

Fourth, reaction trajectories can belay the complex interactions that are taking place in the cell. Results in this chapter show that both metrics of cell behaviour are necessary for an accurate description of the cell.

Quantifying the responses of simulated biological pathways to common configurations of internal cellular organisation

In chapter 5, we demonstrated how the internal organisation of a virtual cell could alter the outcome of two simulated pathways. The internal organisation of

cells varies not only between cells of different types but also within those of the same type. Differences in organisation between cell types are often explained by the cell's role. In cells of the same type, some variation is natural and to be expected. However, irregularities in the size, position and number of some types of organelle are thought to play a role in some non-infectious diseases. In this chapter, we systematically explored the impact of internal cellular organisation by decomposing it into three aspects: Nuclear to Cytoplasm (NC) ratio, nuclear position and distribution of internal cellular volume. Numerical simulations that explored these parameters were run for both the LVD and BMR system. It is found that concentration gradients, reaction trajectories, peak concentrations and times of arrival at the other side of the virtual cell are all dependent on internal organisation in simulated cells.

The results from this chapter can be summarised into four main observations. First, changing the size of centrally positioned nucleus alters the simulated reaction pathways in different ways. As the NC ratio increases, a trend in the peak concentration produced in both reaction systems is found - the response of the systems are not random.

Second, simulated pathways respond to nuclear positioning in different ways. In the LVD system, at equivalent NC ratios, nuclei positioned either centrally, or at the origin or measurement points of the simulated pathway result in higher concentrations than at any other point in the cell. As the NC ratio increases, the differences in these concentrations also increases. In addition, at higher NC ratios, some nuclear positioning can result in very heavy damping of the LVD system, nearly destroying its oscillatory nature. In the BMR system, the exact converse is true with those special positions resulting in lower concentrations than at any other point in the cell and as NC ratio increases the differences between these measurements *decreases*. Further differences between the reaction systems are found when we are presented with concentration gradient results. Repositioning the nucleus dramatically alters wavefront shapes in the LVD system while not modifying concentration gradient to an appreciable degree. In the

BMR system, wavefront shape is relatively unaffected by nuclear position while concentration gradients are altered. On aggregate, we find the LVD system to be more sensitive to nuclear position.

Third, distribution of internal volume changes the outcome of the two simulated pathways in different ways. It is again observed that (i) the LVD system is more sensitive to distribution of volume than the BMR system and (ii) this sensitivity increases in magnitude as the Structure to Cytoplasm (SC) ratio increases. At an SC ratio of 20%, there is little difference between reaction trajectories in either system. At an SC ratio of 50%, different random configurations of internal organisation result in completely different oscillatory behaviour in the LVD system, with damping behaviour dramatically affected. The LVD system reaches an equilibrium faster than the BMR system. There is a small increase in sensitivity within the BMR system when going from 20% to 50% SC although reaction trajectories remain mostly similar.

Finally, provided that the internal structures have impermeable membranes, greater SC (structure to cytoplasm) ratios increase the time it takes for a simulated chemical to cross the cell. This is true when comparing virtual cells with similar internal organisation i.e. all cells possessing randomly positioned nuclei, or all cells possessing distributed internal volume, or all cells having a centrally positioned nuclei. This is the only result that is independent of the choice of reaction system.

From these observations we are able to draw three conclusions. First, it is clear that alone, the reaction system representing a biological pathway is not enough to predict behaviour. This is not a new claim, but we have measured the different sensitivities to spatial organisation in two common motifs for reaction systems in biological cells. In the vast majority of our results in this chapter it is apparent that there is no single rule that will predict the behaviour within both reaction-diffusion systems when subject to a change of one of the spatial variables.

Second, our results suggest that it is the interaction of cell structure and

reaction parameters that dictate the biochemical behaviour of the cell. Cells may tune their reaction performance by the behaviours identified in chapter 2 to account for their internal organisation. In situations where the cell must change its internal organisation, for example when growing or dividing, reaction parameters may be constantly changed in reaction to this.

Third, on a related note, these results imply that cells may adopt different structures to achieve the same moderating behaviour for different pathways. However, it seems more likely that reaction and diffusion parameters are adapted to the geometry of cell to produce the required outcome. In real cells there are of course multiple reaction systems that may all require their behaviour moderating; given that cells do not change shape in-between the execution of every reaction pathway it is possible that the final shape is a compromise.

Quantifying the responses of simulated biological pathways to common processes that alter morphological configuration

In chapter 6, we show through simulation how natural processes that change the morphology of a cell can alter the dynamics of pathways taking place in them. Natural cells rarely maintain a constant shape throughout their life, but instead change from one shape to another in order to best suit their environment. Simulations in this chapter fall into two categories; those who simulate processes where the basic external shape of the cell remains unchanged but the volume changes, and those in which the external shape of the cell changes, but volume is conserved.

Our simulations show that processes of the former type (hypertrophy and atrophy) rarely affect the dynamics of the simulated cell. On the other hand, we show that processes of the later type - that *do* change the shape of the cell - have the potential to change behaviour.

Our results can be synthesised into two observations.

First, in simulations of processes that do not conserve volume (hypertrophy and atrophy within a model *coccus*, *bacillus*, *spirillum* and irregular cell), the

responses of the LVD and BMR reaction systems are different. In the LVD system, both the reaction trajectories and concentration gradients do not respond to atrophy or hypertrophy. These results are different for the BMR system with cell dynamics being affected. As the model cells become larger, peak concentrations decrease, whilst maintaining reaction trajectory shape, and gradients become less pronounced.

Second, in processes that change the shape, of the cell dynamical response depends on the choice of model pathway. When changing the shape of the cell, the signal produced by the LVD system is often damped, producing differences in peak concentration, whilst in most results the gradient of concentration remains similar. In the BMR system, changes in shape result not only in differences in reaction trajectory, but large differences in concentration gradient.

From these observations we can draw three conclusions. First, processes that change the morphology of the cell have the potential to completely change cell dynamics. However, this depends both on the type of morphological change and the reaction system. Reaction systems with feedback mechanisms such as the LVD system are nearly immune to changes in size of cell, as one of the reaction components can make more of itself without other reactants. Concentration conservative reaction systems such as the BMR system do respond to changes in size because of their lack of ability to self regulate.

Second, changing the size of the cell can serve as a control mechanism for cell dynamics. In non-volume conservative processes, changing the morphology of the cell can change the rate of reaction. In the BMR system, smaller reaction vessels result in faster rates of reaction, whilst maintaining the trajectory motif. Along with enzymatic control (discussed in chapter 2), this could allow the cell more control over its destiny. On the other hand, if the cell needs a process to occur consistently no matter its size, for example, a metabolic process in a muscle cell, which may regularly undergo hypertrophy and atrophy, then that cell may have evolved to use a reaction system with a feedback mechanism so that it is robust to that sort of change. With this in mind is seems more

likely that reaction and diffusion parameters have been adapted to the shape of the cell, rather than the other way around. Satisfactorily determining which direction(s) causality between these two flows will require a variety of further work.

Third, changing the shape of the cell can serve as a control mechanism for cell dynamics. In non-volume conserving processes, changing the morphology of the cell can change the pattern of the reaction. Volume sensitive reactions, such as the BMR system, have the motif of their reaction trajectory completely changed when the cell changes shape. In addition, this simple pathway can give rise to complex concentration gradients behaviour when the shape is changed. Oscillatory systems have the potential to be controlled in interesting ways by changing their shape. Filamentation in particular is responsible for a damping effect.

8.2 General biological implications

Biologists are interested in factors that affect cellular function. Traditionally, computational modellers have approached this issue by studying the effects of changes to metabolic networks. Metabolic networks in their turn are controlled by changes to reaction concentrations and constants. These *in their turn* are tuned under the influence of networks. Altering the cell's metabolism is strongly implicated in disease or other organism malfunction. This thesis, on the other hand, shows another way in which changes to cellular function can be regulated, this time indirectly, by changing the morphology of the cell, internally and externally.

In all chapters of this thesis, we have examined how changes in morphology can induce significant changes in cellular dynamics. In particular, we have seen that changes in internal organisation have the potential to make the greatest difference in behaviour. Even morphological features that have less of an effect on pathway dynamics, for example hypertrophy in cells with feedback reaction

systems, could have effects on overall cell behaviour due to the interrelated, complex nature of biochemical networks. As alluded to in the conclusions for individual research questions, there is a strong possibility that metabolic parameters had to be tuned to deal with evolutionary changes in cell morphology. The grounds for this conclusion are further strengthened when we consider, that within a single cell, many thousands of reactions take place, but the cell only has *one* morphology at a particular point in time.

Modelling the effects of morphology upon real cellular pathways could afford us the opportunity to identify specifically how structural changes can lead to disease. A detailed model of a particular cell and reaction system could help identify morphological markers as a diagnostic measure.

As discussed in the introduction, a bacterium's shape is crucial to its metabolic processes. Our work has examined metabolic effects caused by this shape change. Our investigations give biologists a handle on reaction regulation brought about by this behaviour.

8.3 Limitations and Further work

The work presented here explores a new way of studying cellular function, made possible by the recent explosion of computational power made available easily and cheaply. This initial study looked primarily at a few morphological features and explored a few abstractions of both cells and reactions, but we envisage an immediate applicability to more realistic scenarios:

8.3.1 The shape of internal cellular structures has not been modelled

In this thesis we quantified the response of simulated pathways to three aspects of cell morphology: (i) external cell shape (ii) internal cellular organisation and (iii) processes that change the morphology of the cell. These simulations were performed with detailed external cellular shapes, but with abstract representa-

tions for the innards of the cell - which demonstrated the greatest potential for control. Variably sized spheres were used to represent all organelles and other inclusions. However, natural cells, both eurkaryotic and prokaryotic, display an enormous variance in organelles morphology; the fibres that give the cytoplasm its structure, known as the cytoskeleton, are long thin pieces of biological scaffolding, the endoplasmic reticulum takes on a highly "folded" appearance and is studded with small ribosomes and Golgi apparatus. New simulations in this field should further refine our understanding of how internal cellular configuration alters the behaviour of pathways by modelling these organelles as they appear *in vivo*. Such simulations could also start to answer the question of "How much detail is enough?" Do cells need to be modelled as a *true intracellular space* to gain the greatest insights, or are abstractions of internal shape enough?

8.3.2 Varying initial position

The simulations in thesis initially locate pathway chemicals at one side of the cell and measure concentration at the other. Experiments in Chapter 5, in which the initial location of the chemicals is moved to the centre (modelling a signalling pathway that starts at the nucleus and travels out through the cell) show that other starting locations are also sensitive to cell shape. The timeseries of concentration recorded in the *Coccus* is especially dependant on the starting location. While these results are exciting, there are many other realistic initial pathway chemical locations that could be modelled. One example of future work could simulate signalling molecules entering the cell from all directions and propagating toward the nucleus.

8.3.3 The response of a specific cellular system to a range of plausible cell morphologies

This thesis has successfully quantified the response of two abstractions of cellular pathways to a wide range of morphologies. Future simulations might instead

faithfully recreate the dynamics of one particular pathway of interest, and simulate it in a smaller, but more representative, range of morphologies of the host cell. Suitable cells may include those with particularly interesting internal configurations or those that change their shape frequently, such as a neutrophil.

8.3.4 Integrating with cell imaging data

Cells have intricate internal structures, and imaging techniques such as super high resolution imaging and photo-bleaching [Ball et al., 2012], capable of capturing them in enough detail to discern 3D geometry, have the potential to greatly enhance spatial simulations of cells. These techniques are still in an experimental and low throughput stage and consequently no robust methods exist to gather a large amount of empirical statistics on cellular geometry. Once available, this geometric data could be employed to greatly enhance the level of detail in cellular dynamics simulations. Without the empirical data on which to base detailed simulations, it has been argued that the field of cell dynamics simulation is limited [Gibson and Gibson, 2009].

Even when these parameters have been found, the conversion of experimental data into simulation parameters can prove quite challenging in terms of setting up the *in silico* initial conditions: it is difficult to specify three-dimensional shapes and distributions in an intuitive way. One approach is to use image processing algorithms to extract qualitative features from microscopy images and videos at the same time as noise-reduction algorithms are applied. As an example, PhenoPlot [Sailem et al., 2015] enables the automatic, accurate translation of high resolution microscopy images of cells to convenient glyph based representations of the cell in terms of features such as nuclear texture and size, degree of roundedness, etc. These features can then be used to tailor the characteristics of the virtual cell to be simulated using Reaction-Diffusion Cell (ReDi-Cell). Thus, with these combined techniques the simulation of more realistic systems within an integrated experimental/computational cycle becomes feasible.

8.4 Final remarks

While quantitative models of metabolic and genetic networks have provided a lot of insight into cellular function in the last few years, they do not provide a full picture. In this thesis, we provide an in-depth study of the importance of considering morphological features, such as shape of cells and internal configuration. This has been done by quantitative simulations of abstract models, but the results demonstrate that such morphological detail *must* be considered in order to create truly accurate models, even though they increase the computational cost of modeling and simulation. We demonstrate how a reaction-diffusion system can be used alongside a mechanism to specify three-dimensional information, and our use of easily available General Purpose Graphics Processing Unit Computing (GPGPU) points the way to how modeling can be increasingly accurate with currently available resources. In this thesis, we have addressed the lack of precise cellular geometry information by simulating abstract, general geometries. The virtue of performing this type of simulation is that instead of drawing conclusions about specific systems we are able to draw conclusions about how morphology affects cells in *general*. We hope our results encourage the development of better models which can aid in improved understanding of biological processes.

Bibliography

Modelica. `http://blog.xogeny.com/blog/eqs-to-comps-2/`, 2014. Accessed: 2014-5-9.

Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, James D Watson, and AV Grimstone. 1995. Molecular biology of the cell (3rd edn). *Trends in Biochemical Sciences*, 20(5):210–210.

Bruce Alberts, Dennis Bray, Karen Hopkin, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Essential cell biology*. Garland Science, 2013.

Linda JS Allen, Fred Brauer, Pauline Van den Driessche, and Jianhong Wu. *Mathematical epidemiology*. Springer, 2008.

KI Anderson, Yu-Li Wang, and JV Small. Sep 1996. Coordination of protrusion and translocation of the keratocyte involves rolling of the cell body. *J Cell Biol*, 134(5).

Steven S Andrews, Tuan Dinh, and Adam P Arkin. 2009. Stochastic models of biological processes. In *Encyclopedia of Complexity and Systems Science*, pages 8730–8749. Springer.

Stephen Anthony, Liangfang Zhang, and Steve Granick. 2006. Methods to track single-molecule trajectories. *Langmuir*, 22(12):5266–5272.

Adam Arkin, John Ross, and Harley H McAdams. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells. *Genetics*, 149(4):1633–1648.

David K Arrowsmith and Colin M Place. *An introduction to dynamical systems*. Cambridge University Press, 1990.

Kendall Atkinson, Weimin Han, and David E Stewart. *Numerical solution of ordinary differential equations*, volume 108. John Wiley & Sons, 2011.

Daniel Axelrod, DE Koppel, J Schlessinger, Ei Elson, and WW Webb. 1976. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophysical journal*, 16(9):1055.

Graeme Ball, Richard M Parton, Russell S Hamilton, and Ilan Davis. 2012. A cell biologist's guide to high resolution imaging. *Meth. Enzymol.*, 504:29–55.

Philip Ball. 2015. Forging patterns and making waves from biology to geology: a commentary on turing (1952)the chemical basis of morphogenesis. *Phil. Trans. R. Soc. B*, 370(1666):20140218.

TJ Beveridge. 1988. The bacterial surface: general considerations towards design and function. *Canadian Journal of Microbiology*, 34(4):363–372.

Nick Bisceglia. Cell Biology. `http://www.nature.com/scitable/topic/cell-biology-13906536`, 2016. Accessed: 2016-8-1.

James L Bischoff and David Z Piper. *Marine geology and oceanography of the Pacific manganese nodule province*. Springer Science & Business Media, 2013.

MA Bokhari and IS Sadek. 2011. Optimal control problem of reaction-diffusion economic system. In *Modeling, Simulation and Applied Optimization (ICMSAO), 2011 4th International Conference on*, pages 1–5. IEEE.

Fred Brauer, Carlos Castillo-Chavez, and Carlos Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 40. Springer, 2001.

Paul C Bressloff. *Stochastic processes in cell biology*, volume 41. Springer, 2014.

Michael Brin and Garrett Stuck. *Introduction to dynamical systems*. Cambridge University Press, 2002.

Henk Broer and Floris Takens. *Dynamical systems and chaos*, volume 172. Springer Science & Business Media, 2010.

Joshua J. Buchman and Li-Huei Tsai. 2008. Putting a notch in our understanding of nuclear migration. *Cell*, 134(6):912 – 914. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2008.09.007. URL `http://www.sciencedirect.com/science/article/pii/S0092867408011264`.

Matthew T Cabeen and Christine Jacobs-Wagner. 2005. Bacterial cell shape. *Nature Reviews Microbiology*, 3(8):601–610.

Robert Stephen Cantrell and Chris Cosner. *Spatial ecology via reaction-diffusion equations*. John Wiley & Sons, 2004.

celldynamics.johnshopkins.edu. What are cellular dynamics. `http://celldynamics.johnshopkins.edu`, 2014. Accessed: 2015-1-3.

Luonan Chen, Ruiqi Wang, Chunguang Li, and Kazuyuki Aihara. *Modeling biomolecular networks in cells: structures and dynamics*. Springer Science & Business Media, 2010.

Xi Chen, Liang Guo, Jianhong Kang, Yunlong Huo, Shiqiang Wang, and Wenchang Tan. February 2014. Calcium waves initiating from the anomalous subdiffusive calcium sparks. *J R Soc Interface*, 11(91):20130934–20130934.

Shouqiang Cheng, Yu Liu, Christopher S Crowley, Todd O Yeates, and Thomas A Bobik. 2008. Bacterial microcompartments: their properties and paradoxes. *Bioessays*, 30(11-12):1084–1095.

Hank Childs, Eric Brugger, Brad Whitlock, Jeremy Meredith, Sean Ahern, David Pugmire, Kathleen Biagas, Mark Miller, Cyrus Harrison, Gunther H. Weber, Hari Krishnan, Thomas Fogal, Allen Sanderson, Christoph Garth, E. Wes Bethel, David Camp, Oliver Rübel, Marc Durant, Jean M. Favre, and Paul Navrátil. Oct 2012. VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data. In *High Performance Visualization–Enabling Extreme-Scale Scientific Insight*, pages 357–372.

K. H. Chow, R. E. Factor, and K. S. Ullman. Mar 2012. The nuclear envelope environment and its cancer connections. *Nat. Rev. Cancer*, 12(3):196–209.

C Cosner. 2008. Reaction–diffusion equations and ecological modeling. In *Tutorials in Mathematical Biosciences IV*, pages 77–115. Springer.

Julien Courchet, Tommy L Lewis, Sohyon Lee, Virginie Courchet, Deng-Yuan Liou, Shinichi Aizawa, and Franck Polleux. 2013. Terminal axon branching is regulated by the lkb1-nuak1 kinase pathway via presynaptic mitochondrial capture. *Cell*, 153(7):1510–1525.

Markus W Covert, Christophe H Schilling, Iman Famili, Jeremy S Edwards, Igor I Goryanin, Evgeni Selkov, and Bernhard O Palsson. 2001. Metabolic modeling of microbial strains in silico. *Trends in biochemical sciences*, 26(3): 179–186.

Ann E Cowan, Ion I Moraru, James C Schaff, Boris M Slepchenko, and Leslie M Loew. 2012. Spatial modeling of cell signaling networks. *Methods in cell biology*, 110:195.

John Crank. *The mathematics of diffusion*. Oxford university press, 1979.

Nicoline Y den Breems, Lan K Nguyen, and Don Kulasiri. October 2014. Integrated signaling pathway and gene expression regulatory model to dissect dynamics of Escherichia coli challenged mammary epithelial cells. *BioSystems*, 126C:27–40.

Xavier Dolcet, David Llobet, Judit Pallares, and Xavier Matias-Guiu. 2005. NF-kB in development and progression of human cancer. *Virchows Arch*, 446:475–482.

Gerald W Dorn, Rick B Vega, and Daniel P Kelly. 2015. Mitochondrial biogenesis and dynamics in the developing and diseased heart. *Genes & development*, 29(19):1981–1991.

Isabelle Dupin and Sandrine Etienne-Manneville. 2011. Nuclear positioning: Mechanisms and functions. *The International Journal of Biochemistry and Cell Biology*, 43(12):1698 – 1707. ISSN 1357-2725. doi: http://dx.doi.org/10.1016/j.biocel.2011.09.004. URL `http://www.sciencedirect.com/science/article/pii/S1357272511002548`.

Lisa J. Edens, Karen H. White, Predrag Jevtic, Xiaoyang Li, and Daniel L. Levy. 2013. Nuclear size regulation: from single cells to development and disease. *Trends in Cell Biology*, 23(4):151–159. ISSN 0962-8924. doi: 10.1016/j.tcb.2012.11.004. URL `http://dx.doi.org/10.1016/j.tcb.2012.11.004`.

M Ehrenberg and R Rigler. 1974. Rotational brownian motion and fluorescence intensify fluctuations. *Chemical Physics*, 4(3):390–401.

Narat J Eungdamrong and Ravi Iyengar. December 2004. Computational approaches for modeling regulatory cellular networks. *Trends in cell biology*, 14 (12):661–669.

Richard Eva, Elisa Dassie, Patrick T Caswell, Gunnar Dick, Jim C Norman, James W Fawcett, et al. 2010. Rab11 and its effector rab coupling protein contribute to the trafficking of β1 integrins during axon growth in adult dorsal root ganglion neurons and pc12 cells. *The Journal of Neuroscience*, 30(35): 11654–11669.

Richard Eva, Sarah Crisp, Jamie RK Marland, Jim C Norman, Venkateswarlu Kanamarlapudi, James W Fawcett, et al. 2012. Arf6 directs axon transport and traffic of integrins and regulates axon growth in adult drg neurons. *The Journal of Neuroscience*, 32(30):10352–10364.

David Fell and Athel Cornish-Bowden. *Understanding the control of metabolism*, volume 2. Portland press London, 1997.

Johannes Fels, Sergei N. Orlov, and Ryszard Grygorczyk. May 2009. The hydrogel nature of mammalian cytoplasm contributes to osmosensing and

extracellular ph sensing. *Biophys J*, 96(10):4276–4285. ISSN 0006-3495. doi: 10.1016/j.bpj.2009.02.038.

Adolph Fick. 1855. V. on liquid diffusion. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(63):30–39.

BE Foreman, G Nelson, V See, CA Horton, DG Spiller, SW Edwards, HP McDowell, JF Unitt, E Sullivan, R Grimley, et al. 2004. Oscillations in nf-kappab signaling control the dynamics of gene expression. *Science*, 306(5696):704708.

Leonard J. Foster, Carmen L. de Hoog, Yanling Zhang, Yong Zhang, Xiaohui Xie, Vamsi K. Mootha, and Matthias Mann. 2006. A mammalian organelle map by protein correlation profiling. *Cell*, 125(1):187–199. ISSN 0092-8674. doi: 10.1016/j.cell.2006.03.022. URL `http://dx.doi.org/10.1016/j.cell.2006.03.022`.

Joseph Fourier. *Theorie analytique de la chaleur, par M. Fourier*. Chez Firmin Didot, père et fils, 1822.

Sonke Friedrichsen, Claire V Harper, Sabrina Semprini, Michael Wilding, Antony D Adamson, Dave G Spiller, Glyn Nelson, John J Mullins, Michael RH White, and Julian RE Davis. 2006. Tumor necrosis factor-α activates the human prolactin gene promoter via nuclear factor-κb signaling. *Endocrinology*, 147(2):773–781.

Richard E. Frye and Daniel A. Rossignol. May 2011. Mitochondrial dysfunction can connect the diverse medical symptoms associated with autism spectrum disorders. *Pediatr Res*, 69(5 Pt 2):41R–47R. ISSN 0031-3998. doi: 10.1203/PDR.0b013e318212f16b. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3179978/`.

Brian G Fuller, Michael A Lampson, Emily A Foley, Sara Rosasco-Nitcher, Kim V Le, Page Tobelmann, David L Brautigan, P Todd Stukenberg, and Tarun M Kapoor. June 2008. Midzone activation of aurora B in anaphase

produces an intracellular phosphorylation gradient. *Nature*, 453(7198):1132–1136.

Michael A Gibson and Jehoshua Bruck. 2000. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889.

William T Gibson and C Gibson, Matthew. *Current Topics in Developmental Biology*. Elsivier, 2009.

Daniel T Gillespie. December 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361.

Daniel T Gillespie. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434.

Daniel T Gillespie. 1992. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1):404–425.

Daniel T Gillespie. 2000. The chemical langevin equation. *The Journal of Chemical Physics*, 113(1):297–306.

Daniel T Gillespie. 2001. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733.

Roland Glowinski, Stanley J Osher, and Wotao Yin. *Splitting Methods in Communication, Imaging, Science, and Engineering*. Springer, 2017.

Magdalena R Golachowska, Dick Hoekstra, and Sven CD van IJzendoorn. 2010. Recycling endosomes in apical plasma membrane domain formation and epithelial cell polarity. *Trends in cell biology*, 20(10):618–626.

Haijun Gong, Paolo Zuliani, Anvesh Komuravelli, James R Faeder, and Edmund M Clarke. 2010. Analysis and verification of the HMGB1 signaling pathway. *BMC Bioinformatics*, 11 Suppl 7(Suppl 7):S10.

KE Gordon, IMM Van Leeuwen, S Lain, and MAJ Chaplain. 2009. Spatio-temporal modelling of the p53–mdm2 oscillatory system. *Mathematical Modelling of Natural Phenomena*, 4(3):97–116.

Michael Gösch and Rudolf Rigler. 2005. Fluorescence correlation spectroscopy of molecular motions and kinetics. *Advanced drug delivery reviews*, 57(1): 169–190.

Douglas R Green. 1998. Apoptotic pathways: the roads to ruin. *Cell*, 94(6): 695–698.

Gregg G. Gundersen and Howard J. Worman. 2013. Nuclear positioning. *Cell*, 152(6):1376 – 1389. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2013.02.031. URL `http://www.sciencedirect.com/science/article/pii/S0092867413002201`.

Hong-Sup Hahn, Peter J Ortoleva, and John Ross. 1973. Chemical oscillations and multiple steady states due to variable boundary permeability. *Journal of theoretical biology*, 41(3):503–521.

Alan Hastings. 1978. Global stability in lotka-volterra systems with diffusion. *Journal of Mathematical Biology*, 6(2):163–168.

Elke Haustein and Petra Schwille. 2004. Single-molecule spectroscopic methods. *Current opinion in structural biology*, 14(5):531–540.

Reinhart Heinrich and Stefan Schuster. *The regulation of cellular systems*. Springer Science & Business Media, 2012.

B Hess and A Boiteux. 1971. Oscillatory phenomena in biochemistry. *Annual Review of Biochemistry*, 40(1):237–258.

B Hess, A Boiteux, B Chance, EK Pye, AK Ghosh, and B Hess. 1973. Biological and biochemical oscillators. *Substrate Control of Glycolytic Oscillations, Academic Press, New York*, pages 229–241.

Desmond J Higham. 2001. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, 43(3):525–546.

Desmond J Higham. 2008. Modeling and simulating chemical reactions. *SIAM review*, 50(2):347–368.

Yujiro Higuchi, Peter Ashwin, Yvonne Roger, and Gero Steinberg. 2014. Early endosome motility spatially organizes polysome distribution. *The Journal of cell biology*, 204(3):343–357.

William R Holmes and Leah Edelstein-Keshet. 2012. A comparison of computational models for eukaryotic cell shape and motility. *PLoS Comput Biol*, 8 (12):e1002793.

Michael Hoppert and Frank Mayer. 1999. Principles of macromolecular organization and cell function in bacteria and archaea. *Cell Biochemistry and Biophysics*, 31(3):247–284. ISSN 1559-0283. doi: 10.1007/BF02738242. URL `http://dx.doi.org/10.1007/BF02738242`.

Thomas R Horton. 2006. The number of nuclei in basidiospores of 63 species of ectomycorrhizal homobasidiomycetes. *Mycologia*, 98(2):233–238.

Martin Howard. September 2006. Cell signalling: changing shape changes the signal. *Curr. Biol.*, 16(17):R673–5.

Philip Hugenholtz. 2002. Exploring prokaryotic diversity in the genomic era. *Genome biology*, 3(2):1.

Yongyun Hwang, Praveen Kumar, and Abdul I Barakat. July 2014. Intracellular regulation of cell signaling cascades: how location makes a difference. *Journal of mathematical biology*, 69(1):213–242.

AEC Ihekwaba, DS Broomhead, RL Grimley, N Benson, and DB Kell. 2004. Sensitivity analysis of parameters controlling oscillatory signalling in the nf-kb pathway: the roles of ikk and ikba. *Syst Biol*, 1:93–103.

Silvana Ilie, Wayne H Enright, and Kenneth R Jackson. 2009. Numerical solution of stochastic models of biochemical kinetics. *Canadian Applied Mathematics Quarterly*, 17(3):523–554.

Jun-ichiro Inoue, Jin Gohda, Taishin Akiyama, and Kentaro Semba. 2007. Nf-κb activation in development and progression of cancer. *Cancer science*, 98 (3):268–274.

Michael Charles Irwin. *Smooth dynamical systems*, volume 17. World Scientific, 2001.

Jürgen Jost. *Dynamical systems: examples of complex behaviour*. Springer Science & Business Media, 2006.

Mario Leon Juncosa and David Young. 1957. On the crank-nicolson procedure for solving parabolic partial differential equations. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 53, pages 448–461. Cambridge Univ Press.

Gerald Karp. *Cell biology*. Wiley Online Library, 1979.

J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, and M. W. Covert. July 2012. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150: 389–401.

Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54. Cambridge university press, 1997.

M Katsoulis, J Lekka, I Vlachonikolis, and GS Delides. 1995. The prognostic value of morphometry in advanced epithelial ovarian cancers. *British journal of cancer*, 72(4):958.

Kenneth J Kauffman, Purusharth Prakash, and Jeremy S Edwards. 2003. Advances in flux balance analysis. *Current opinion in biotechnology*, 14(5):491–496.

Cheryl A Kerfeld, Sabine Heinhorst, and Gordon C Cannon. 2010. Bacterial microcompartments. *Microbiology*, 64(1):391.

Boris N Kholodenko. March 2006. Cell-signalling dynamics in time and space. *Nature reviews. Molecular cell biology*, 7(3):165–176.

Boris N Kholodenko. December 2009. Spatially distributed cell signalling. *FEBS letters*, 583(24):4006–4012.

Boris N Kholodenko and Walter Kolch. May 2008. Giving space to cell signaling. *Cell*, 133(4):566–567.

Hong Jin Kim, N Hawke, and Albert S Baldwin. 2006. Nf-κb and ikk as therapeutic targets in cancer. *Cell Death & Differentiation*, 13(5):738–747.

Peter E Kloeden and Eckhard Platen. 1992. Higher-order implicit strong numerical schemes for stochastic differential equations. *Journal of statistical physics*, 66(1-2):283–314.

Omar M Knio, Habib N Najm, and Peter S Wyckoff. 1999. A semi-implicit numerical scheme for reacting flow: Ii. stiff, operator-split formulation. *Journal of Computational Physics*, 154(2):428–467.

Shigeru Kondo and Takashi Miura. 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. *science*, 329(5999):1616–1620.

Oleg Krichevsky and Grégoire Bonnet. 2002. Fluorescence correlation spectroscopy: the technique and its applications. *Reports on Progress in Physics*, 65(2):251.

Sandeep Krishna, Mogens H Jensen, and Kim Sneppen. 2006. Minimal model of spiky oscillations in nf-κb signaling. *Proceedings of the National Academy of Sciences*, 103(29):10840–10845.

Ekat Kritikou. 2008. Cell morphology: Breaking the spatial code. *Nature Reviews Molecular Cell Biology*, 9(7):500–500.

Ulrich Kubitscheck, Oliver Kückmann, Thorsten Kues, and Reiner Peters. 2000. Imaging and tracking of single gfp molecules in solution. *Biophysical Journal*, 78(4):2170–2179.

Hao Yuan Kueh and Timothy J Mitchison. August 2009. Structural plasticity in actin and tubulin polymer dynamics. *Science*, 325(5943):960–963.

Fred M Kusumoto. *Cardiovascular pathophysiology*. Hayes Barton Press, 2004.

Edward J Lesnefsky, Shadi Moghaddas, Bernard Tandler, Janos Kerner, and Charles L Hoppel. 2001. Mitochondrial dysfunction in cardiac disease: ischemia–reperfusion, aging, and heart failure. *Journal of molecular and cellular cardiology*, 33(6):1065–1089.

Anthony W Leung. 1988. Diffusion-reaction systems in neutron-fission reactors and ecology. In *Nonlinear Diffusion Equations and Their Equilibrium States II*, pages 41–53. Springer.

Hui Li, Shuo-Xing Dou, Yu-Ru Liu, Wei Li, Ping Xie, Wei-Chi Wang, and Peng-Ye Wang. 2015. Mapping intracellular diffusion distribution using single quantum dot tracking: compartmentalized diffusion defined by endoplasmic reticulum. *Journal of the American Chemical Society*, 137(1):436–444.

L M Loew and J C Schaff. October 2001. The Virtual Cell: a software environment for computational cell biology. *Trends in biotechnology*, 19(10): 401–406.

A J Lotka. Elements of physical biology, 1925.

Vigirdas Mackevicius. *Introduction to stochastic analysis: integrals and differential equations*. John Wiley & Sons, 2013.

Michael T Madigan, David P Clark, David Stahl, and John M Martinko. *Brock Biology of Microorganisms 13th edition*. Benjamin Cummings, 2010.

Celine I Maeder, Mark A Hink, Ali Kinkhabwala, Reinhard Mayr, Philippe I H Bastiaens, and Michael Knop. November 2007. Spatial regulation of

Fus3 MAP kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling. *Nature cell biology*, 9(11):1319–1326.

Douglas Magde, Elliot L Elson, and Watt W Webb. 1974. Fluorescence correlation spectroscopy. ii. an experimental realization. *Biopolymers*, 13(1):29–61.

Larry L Mai, M Young Owl, and M Patricia Kersting. *The Cambridge Dictionary of Human Biology and Evolution.* Cambridge University Press, 2005.

PhilipáK Maini, KevináJ Painter, and HeleneáNguyen Phong Chau. 1997. Spatial pattern formation in chemical and biological systems. *Journal of the Chemical Society, Faraday Transactions*, 93(20):3601–3610.

Elizabeth Martin. *Dictionary of life sciences.* Pica Press: distributed by Universe Books, 1977.

Tomáš Mazel, Rebecca Raymond, Mary Raymond-Stintz, Stephen Jett, and Bridget S Wilson. 2009. Stochastic modeling of calcium in 3d geometry. *Biophysical journal*, 96(5):1691–1706.

Heidi M McBride, Margaret Neuspiel, and Sylwia Wasiak. 2006. Mitochondria: more than just a powerhouse. *Current Biology*, 16(14):R551–R560.

John F. McLaughlin and Jonathan Roughgarden. 1991. Pattern and stability in predator-prey communities: How diffusion in spatially variable environments affects the lotka-volterra model. *Theoretical Population Biology*, 40(2):148 – 172. ISSN 0040-5809. doi: http://dx.doi.org/10.1016/0040-5809(91)90051-G.

Shawn Means, Alexander J Smith, Jason Shepherd, John Shadid, John Fowler, Richard JH Wojcikiewicz, Tomas Mazel, Gregory D Smith, and Bridget S Wilson. 2006. Reaction diffusion modeling of calcium dynamics with realistic er geometry. *Biophysical journal*, 91(2):537–557.

Jason Meyers, Jennifer Craig, and David J Odde. September 2006. Potential for control of signaling pathways via cell size and shape. *Current biology : CB*, 16(17):1685–1693.

R.A. Meyers. *Epigenetic Regulation and Epigenomics: Advances in Molecular Biology and Medicine*. Wiley-Blackwell, 2012. ISBN 9783527326822. URL `https://books.google.co.uk/books?id=jzU7wpL0I40C`.

Ron Milo and Rob Philips. *Cell Biology by the Numbers*. Garland Science, 2016.

GN Mil'shtejn. 1975. Approximate integration of stochastic differential equations. *Theory of Probability & Its Applications*, 19(3):557–562.

R Montironi, M Scarpelli, S Sisti, A Braccischi, and GM Mariuzzi. 1990. Morphological and quantitative analyses of intraductal dysplasia of the prostate. *Analytical cellular pathology: the journal of the European Society for Analytical Cellular Pathology*, 2(5):277–285.

Britney L Moss, Shimon Gross, Seth T Gammon, Anant Vinjamoori, and David Piwnica-Worms. 2008. Identification of a ligand-induced transient refractory period in nuclear factor-κb signaling. *Journal of Biological Chemistry*, 283 (13):8687–8698.

Brian Munsky and Mustafa Khammash. 2006. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of chemical physics*, 124(4):044104.

myphysicslab.com. A derivation of multi-variable Runge-Kutta 4. `http://www.myphysicslab.com`, 2015. Accessed: 2015-5-9.

David L Nelson, Albert L Lehninger, and Michael M Cox. *Lehninger principles of biochemistry*. Macmillan, 2008.

Susana R Neves and Ravi Iyengar. February 2009. Models of spatially restricted biochemical reaction systems. *J. Biol. Chem.*, 284(9):5445–5449.

Susana R Neves, Panayiotis Tsokas, Anamika Sarkar, Elizabeth A Grace, Padmini Rangamani, Stephen M Taubenfeld, Cristina M Alberini, James C

Schaff, Robert D Blitzer, Ion I Moraru, et al. 2008. Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks. *Cell*, 133(4):666–680.

Philipp Niethammer, Philippe Bastiaens, and Eric Karsenti. March 2004. Stathmin-tubulin interaction gradients in motile and mitotic cells. *Science (New York, N.Y.)*, 303(5665):1862–1866.

Béla Novák and John J Tyson. 2008. Design principles of biochemical oscillators. *Nature reviews Molecular cell biology*, 9(12):981–991.

RM Noyes and RJ Field. 1974. Oscillatory chemical reactions. *Annual Review of Physical Chemistry*, 25(1):95–119.

Clare M O'Connor, Jill U Adams, and Jennifer Fairman. *Essentials of cell biology*. Cambridge: NPG Education, 2010.

Daisuke Ohshima, Jun-ichiro Inoue, and Kazuhisa Ichikawa. 2012. Roles of spatial parameters on the oscillation of nuclear nf-κb: computer simulations of a 3d spherical cell. *PloS one*, 7(10):e46911.

Richard J Orton, Oliver E Sturm, Vladislav Vyshemirsky, Muffy Calder, David R Gilbert, and Walter Kolch. December 2005. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem. J.*, 392(Pt 2):249–261. doi: 10.1042/BJ20050908.

Hans G. Othmer. *Cell-Based, Continuum and Hybrid Models of Tissue Dynamics*. Springer International Publishing, 2016.

Edward Ott. *Chaos in dynamical systems*. Cambridge university press, 2002.

Brian Owens. Apr 2014. Cell physiology: The changing colour of fat. *Nature*, 508(7496):S52–S53. ISSN 0028-0836. URL `http://dx.doi.org/10.1038/508S52a`. Outlook.

Julius Parulek, Milo S Sramek, Michal Cerve ansky, Marta Novotova, and Ivan Zahradni. 2009. A cell architecture modeling system based on quantitative ultrastructural characteristics. *Systems Biology*, pages 289–312.

Reiner Peters, Jutta Peters, Karl Heinz Tews, and Wolfgang Bähr. 1974. A microfluorimetric study of translational diffusion in erythrocyte membranes. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 367(3):282–294.

Tuan D Pham and Kazuhisa Ichikawa. 2013. Spatial chaos and complexity in the intracellular space of cancer and normal cells. *Theoretical Biology and Medical Modelling*, 10(1):1.

Kenneth J Pienta and Donald S Coffey. 1991. Correlation of nuclear morphometry with progression of breast cancer. *Cancer*, 68(9):2012–2016.

W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. *Numerical Methods in C++*. Cambridge University Press; 3 edition, 2002.

Michael S Rappé, Stephanie A Connon, Kevin L Vergin, and Stephen J Giovannoni. 2002. Cultivation of the ubiquitous sar11 marine bacterioplankton clade. *Nature*, 418(6898):630–633.

Dean Rickles, Penelope Hawe, and Alan Shiell. Nov 2007. A simple guide to chaos and complexity. *J Epidemiol Community Health*, 61(11):933–937. ISSN 0143-005X. doi: 10.1136/jech.2006.054254. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2465602/`.

Lukasz Sadowski, Iwona Pilecka, and Marta Miaczynska. 2009. Signaling from endosomes: location makes a difference. *Experimental cell research*, 315(9): 1601–1609.

Heba Z Sailem, Julia E Sero, and Chris Bakal. 2015. Visualizing cellular imaging data using PhenoPlot. *Nat Commun*, 6:5825.

Isabelle Salmon, Robert Kiss, B Franc, Plínio Gasperin, R Heimann, Jean Lambert Pasteels, and Alain Verhest. 1992. Comparison of morphonuclear features

in normal, benign and neoplastic thyroid tissue by digital cell image analysis. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*, 14(1):47–54.

Allen R Sanderson, Robert M Kirby, Chris R Johnson, and Lingfa Yang. 2006. Advanced reaction-diffusion models for texture synthesis. *Journal of Graphics, GPU, and Game Tools*, 11(3):47–71.

Faiz Sayyid and Sara Kalvala. 2016. On the importance of modelling the internal spatial dynamics of biological cells. *Journal of Biosystems*, 145:53 – 66. ISSN 0303-2647. doi: http://dx.doi.org/10.1016/j.biosystems.2016.05.012.

Ivo F Sbalzarini and Petros Koumoutsakos. 2005. Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of structural biology*, 151(2):182–195.

Heide N Schulz and Bo Barker Jørgensen. 2001. Big bacteria. *Annual Reviews in Microbiology*, 55(1):105–137.

Jagesh V Shah. 2010. Cells in tight spaces: the role of cell shape in cell function. *The Journal of cell biology*, 191(2):233–236.

Zu-Hang Sheng and Qian Cai. 2012. Mitochondrial transport in neurons: impact on synaptic homeostasis and neurodegeneration. *Nature Reviews Neuroscience*, 13(2):77–93.

Martin B Short, P Jeffrey Brantingham, Andrea L Bertozzi, and George E Tita. 2010. Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences*, 107(9):3961–3965.

Philip Siekevitz. 1957. Powerhouse of the cell. *Scientific American*.

G D Smith. *Numerical solutions of partial differntial equations: finite difference methods*. OUP Oxford; 3rd Edition, 1985.

Siowling Soh, Marta Byrska, Kristiana Kandere-Grzybowska, and Bartosz A Grzybowski. June 2010. Reaction-diffusion systems in intracellular molecular transport and control. *Angew. Chem. Int. Ed. Engl.*, 49(25):4170–4198.

Mirela Spillane, Andrea Ketschek, Tanuja T Merianda, Jeffery L Twiss, and Gianluca Gallo. 2013. Mitochondria coordinate sites of axon branching through localized intra-axonal protein synthesis. *Cell reports*, 5(6):1564–1575.

P. A. Srere. July 2000. Macromolecular interactions: tracing the roots. *Trends Biochem Sc*, 25:150–153.

Marc Sturrock, Alan J Terry, Dimitris P Xirodimas, Alastair M Thompson, and Mark AJ Chaplain. 2011a. Spatio-temporal modelling of the hes1 and p53-mdm2 intracellular signalling pathways. *Journal of Theoretical Biology*, 273(1):15–31.

Marc Sturrock, Alan J Terry, Dimitris P Xirodimas, Alastair M Thompson, and Mark AJ Chaplain. 2011b. Spatio-temporal modelling of the hes1 and p53-mdm2 intracellular signalling pathways. *Journal of Theoretical Biology*, 273(1):15–31.

Alan J Terry. 2014. A minimal spatio-temporal model of the nf-κb signalling pathway exhibits a range of behaviours. *Bulletin of mathematical biology*, 76 (10):2363–2388.

Alan J Terry and Mark AJ Chaplain. 2011. Spatio-temporal modelling of the nf-κb intracellular signalling pathway: the roles of diffusion, active transport, and cell geometry. *Journal of Theoretical Biology*, 290:7–26.

Alan J Terry, Marc Sturrock, J Kim Dale, Miguel Maroto, and Mark AJ Chaplain. 2011. A spatio-temporal model of notch signalling in the zebrafish segmentation clock: conditions for synchronised oscillatory dynamics. *PLoS One*, 6(2):e16980.

Nancy L Thompson, Alena M Lieto, and Noah W Allen. 2002. Recent advances in fluorescence correlation spectroscopy. *Current opinion in structural biology*, 12(5):634–641.

Kenneth Todar. *Todar's online textbook of bacteriology*. University of Wisconsin-Madison Department of Bacteriology Madison, Wis, USA, 2006.

Masaru Tomita. 2001. Whole-cell simulation: a grand challenge of the 21st century. *Trends in Biotechnology*, 19(6):205–210. doi: 10.1016/S0167-7799(01)01636-5. URL `http://dx.doi.org/10.1016/S0167-7799(01)01636-5`.

M.L. Turgeon. *Clinical Hematology: Theory and Procedures*. Clinical Hematology: Theory and Procedures. Lippincott Williams & Wilkins, 2005. ISBN 9780781750073. URL `https://books.google.co.uk/books?id=cHAjsUgegpQC`.

A.M. Turing. 1990. The chemical basis of morphogenesis. *Bulletin of Mathematical Biology*, 52(1-2):153–197. ISSN 0092-8240. doi: 10.1007/BF02459572.

J Tyson and S Kauffman. 1975. Control of mitosis by a continuous biochemical oscillation: synchronization; spatially inhomogeneous oscillations. *Journal of Mathematical Biology*, 1(4):289–310.

John J Tyson. 1991. Modeling the cell division cycle: cdc2 and cyclin interactions. *Proceedings of the National Academy of Sciences*, 88(16):7328–7332.

John J Tyson, Katherine C Chen, and Bela Novak. 2003. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current opinion in cell biology*, 15(2):221–231.

Petra van Bergeijk, Max Adrian, Casper C Hoogenraad, and Lukas C Kapitein. 2015. Optogenetic control of organelle transport and positioning. *Nature*.

Jeroen S van Zon, David K Lubensky, Pim RH Altena, and Pieter Rein ten Wolde. 2007. An allosteric model of circadian kaic phosphorylation. *Proceedings of the National Academy of Sciences*, 104(18):7420–7425.

Rodica Vîlcu and Daniela Bala. 2004. Models of oscillating chemical reactions. particularities of some proposed models for the oscillations. *Analele Universitatii din Bucuresti*.

Donald Voet, Judith G Voet, and Charlotte W Pratt. *Fundamentals of biochemistry: life at the molecular level*. John Wiley and Sons, 2016.

V Volterra. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi, 1927.

Zhi-Cheng Wang, Wan-Tong Li, and Shigui Ruan. 2006. Travelling wave fronts in reaction–diffusion systems with spatio-temporal delays. *Journal of Differential Equations*, 222(1):185–232.

Robert Weinberg. *The biology of cancer*. Garland science, 2013.

Hans Wilhelmsson and Enzo Lazzaro. *Reaction-diffusion problems in the physics of hot plasmas*. CRC Press, 2000.

Maxwell Myer Wintrobe and John P Greer. *Wintrobe's clinical hematology*, volume 1. Lippincott Williams & Wilkins, 2009.

Jun Wu, Pontus Boström, Lauren M Sparks, Li Ye, Jang Hyun Choi, An-Hoa Giang, Melin Khandekar, Kirsi A Virtanen, Pirjo Nuutila, Gert Schaart, Kexin Huang, Hua Tu, Wouter D van Marken Lichtenbelt, Joris Hoeks, Sven Enerbäck, Patrick Schrauwen, and Bruce M Spiegelman. 2012. Beige adipocytes are a distinct type of thermogenic fat cell in mouse and human. *Cell*, 150(2):366–376. ISSN 0092-8674. doi: 10.1016/j.cell.2012.05.016. URL `http://dx.doi.org/10.1016/j.cell.2012.05.016`.

Todd O Yeates, Christopher S Crowley, and Shiho Tanaka. 2010. Bacterial microcompartment organelles: protein shell structure and evolution. *Annual review of biophysics*, 39:185.

Zheng Yin, Heba Sailem, Julia Sero, Rico Ardy, Stephen T C Wong, and Chris

Bakal. December 2014. How cells explore shape space: a quantitative statistical perspective of cellular morphogenesis. *Bioessays*, 36(12):1195–1203.

Kevin D Young. September 2006. The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.*, 70(3):660–703.

Kevin D Young. December 2007. Bacterial morphology: why have different shapes? *Curr. Opin. Microbiol.*, 10(6):596–600.

Robert M Youngson. *Collins dictionary of human biology*. Collins, 2006.

Linda Amaral Zettler, Mitchell L Sogin, and David A Caron. 1997. Phylogenetic relationships between the acantharea and the polycystinea: a molecular perspective on haeckels radiolaria. *Proceedings of the National Academy of Sciences*, 94(21):11411–11416.

Qi Zhao, Ming Yi, and Yan Liu. October 2011. Spatial distribution and dose-response relationship for different operation modes in a reaction-diffusion model of the MAPK cascade. *Physical biology*, 8(5):055004.

8.5 ReDi-Cell

This section outlines the structure of the simulation, ReDi-Cell, used in this thesis to generate results in Chapters 5, 6 and 7.

8.5.1 Simulation input

The simulation environment generation begins with parameters describing the total size of the simulation environment, number of chemicals, initial values of concentration, initial concentration location, spatial step size and timestep size. The size information is used to generate the geometry matrix.

The geometry matrix

The geometry matrix is a 3D matrix that describes the virtual environment inside the simulation. The environment external to the cell, and the cell itself are both represented. ReDi-Cell is able to approximate three different primitive cell shapes. These shapes are generated in 3D space, from equations describing geometrical primitives. In situations where one primitive will not suffice composites are formed. For example, *Bacilli* are best approximated by the mathematical definition of a capsule; they can be built as a composite of a cylinder and two hemispheres. This geometry information is stored in the geometry matrix. Each location is given one of two values; empty space (in which chemical can exist) or organelle space (into which chemicals *may* enter/exist). These values inform the boundary conditions in the simulation and the writing of visualisation output files.

The diffusivity matrix

The diffusivity matrix is a 3D matrix that describes diffusivity at all points within the simulation environment. The diffusivity matrix is the same size as the geometry matrix. Each location within this array contains another array, the values of which are the diffusivity constants for each chemical species at that location in the simulation environment. Values within the array are informed by the geometry matrix. This structure allows different chemicals to move through the simulation environment at different rates. For example, an organelle might permit one chemical to move through it slowly, whilst being impermeable to another.

The concentration matrix

The concentration matrix is a 3D matrix that describes the concentration of chemicals at all points within the simulation environment. The concentration matrix is the same size as the geometry matrix. Each location within this array contains another array, the values of which are the values of concentration for each chemical species at that location in the simulation environment. Values are set by the user at the start of simulation.

8.5.2 Computation

GP-GPU and CUDA

The GPU is a specialised processor that was originally designed to render 3D graphics in real time. Rather than containing a few powerful cores, the GPU is a highly parallel multi-core processor containing very many light weight cores. This architecture is known as Single Instructure Multiple Data (SIMD). SIMD processors are well suited to processing large chunks of data if each small, individual piece of data has the same computation applied to it. This new method of computation is known as General Purpose Graphics Processing Unit Computing (GPGPU).

The CUDA language is the Nvidia corporation's API for performing GPGPU upon Nvidia GPUs. In an ideal use scenario, data is transferred from the host memory to the device (the GPU). Then, many threads are run, each processing a small piece of the data, ideally using a lightweight algorithm that does not exhibit branching. Results are stored in the device memory and then transferred back to the host memory. Transfer speeds are slow and so attempts are made to limit them whenever possible. When the amount of data is so large that it cannot fit in device memory additional transfers are inevitable.

The CUDA API is processed according to C++ syntax rules. CUDA code is separate from the main program code. It is compiled to a compute Kernel that is called from the main program at the appropriate time.

The number of threads that can be run at once is limited by the number of multiprocessing units a GPU possesses; each multiprocessor can run 768 threads at once. That means that a GPU with 28 multiprocessors can run $768 * 28 = 21504$ threads at once. Any addition threads that are scheduled will need to wait. Threads of execution are organised into blocks. A single block is executed by a single multiprocessor. Blocks can be organised in 1,2 or 3 dimensions depending on the nature of the problem. It doesn't matter how the threads are organised, 768 is the maximum number of threads that can be run. If more threads than 21504 are required then then more blocks are needed. If this is the case then those blocks will have to wait to be executed. Once running on the device each thread uses its index in the grid of blocks and threads to determine which part of the data it should process.

Main simulation loop

In this section we describe the algorithm that launches the compute kernel. During the simulation, data is written to the disk for offline analysis. Transferring data to and from the GPU is expensive and thus writing to disk every time step can degrade performance. Therefore we compromise, transferring data and writing to the disk at configurable intervals.

The execution of time steps is "batched" together in configurable output intervals. Execution is paused between batches whilst results are written to the disk. The larger the batch size the faster the simulation will run overall. Time step batch sizes of 500 steps are commonly used in this thesis, which provides a good trade-off between speed and output granularity.

The program has two stages: The main simulation loop, that launches the kernel, and the kernel itself. The kernel launcher runs on the CPU, it divides the simulation into batches and writes output to the disk. The kernel runs on the GPU and is responsible for executing the reaction process in parallel and then the diffusion processes in parallel across all of the GPU cores.

Here we show the main simulation loop:

```
macro timesteps = total simulation length/output interval;
while macro timesteps have not been completed do
    Launch Kernel;
    Write visualisation file using concentration and geometry matrices;
    Write concentration timeseries;
end
```

Algorithm 1: The main simulation loop

The Kernel

The Kernel runs on the GPU and is where the reaction-diffusion algorithm is run. A single timestep within the kernel consists of a diffusion process and a reaction process. First, the ODEs representing the reactions are solved. This process occurs over each part of the cell in which diffusion is allowed to take place. Then the diffusion equation is solved for each reagent, again in the appropriate parts of the cell. This process is known as operator splitting. This technique works by separating the original equation into its constituent terms over a single timestep, computing the solutions to each term, and then combining the individual solutions to form a solution to the original equation. This technique is commonly encountered when solving convection-diffusion PDES, where convection and diffusion are solved separately [Glowinski et al., 2017]. Many reaction or diffusion events are run at the same time, instead of one after the other In a single time step, all reaction events are run after all diffusion events have taken place.

Identify thread index;
Copy geometry matrix to GPU;
Copy concentration matrix to GPU;
Copy diffusivity matrix to GPU;
while *micro timesteps have not been completed* **do**
    Perform reactions on old concentration matrix and store in new concentration matrix;
    Perform diffusion on old concentration matrix and store in new concentration matrix;
    Copy new concentration matrix to old concentration matrix;
    Append timeseries to buffer;
end
Copy concentration matrix back to host;
Copy timeseries buffer back to host;

Algorithm 2: The reaction-diffusion kernel

8.5.3 Output

ReDi-Cell outputs visualisation files to be viewed in VisIt. Cell geometry and concentration data are written to the same file. ReDi-Cell also outputs time-series of concentration in CSV format. These are plotted using MATLAB.

8.5.4 performance

In terms of speed, we find that ReDi-Cell can perform approximately 24x faster than VCell when running a Bi-Molecular Reaction (BMR) simulation inside a *coccus*.

8.6 Choice of Diffusion Coefficients

Molecule	**Medium**	**Diffusion Coefficient** $(\mu m^2/s)$
H^+	Water	7000
H^2O	Nucleus of Chicken Erythrocyte	200
protein ($\approx$ 30kDa GFP)	eukaryotic cell (CHO) cytoplasm	30
protein ($\approx$ 30kDa)	*E. coli* cytoplasm	7-8
morphogen (bicod-GFP)	*D. melanogaster* embryo cytoplasm	7
protein ($\approx$ 40kDa)	*E. coli* cytoplasm	2-4
protein ($\approx$ 70-250 kDa)	*E. coli* cytoplasm	0.4-0.2
protein ($\approx$ 140 kDa Tar-YFP)	*E. coli* cytoplasm	0.2
protein ($\approx$ 70 kDa LacY-YFP)	*E. coli* cytoplasm	0.03
mRNA	various locations	0.005-0.001

Table 8.1: Table of diffusion coefficients representative of cellular environments.
Reproduced from [Milo and Philips, 2016]

The value of the diffusion coefficient for a chemical species can dramatically alter the behaviour of the reaction-diffusion system associated with it. Despite the difficulty in measuring diffusion coefficients *some* values of biologically relevant entities have been recorded in the literature; a compilation of common cellular chemicals with their diffusion coefficients can be found in [Milo and Philips, 2016]. Cellular diffusion coefficients are proportional to both the mass of the species and to the density of the medium through which they travel, resulting in a wide range of values. For example in the data in [Milo and Philips, 2016] diffusion coefficients range from $7000\mu m^2/s$ for H^+ ions to $0.001\mu m^2/s$ for mRNA. Diffusion coefficients for protein translocation within the cytoplasm have been measured in *E. coli* to be between $30\mu m^2/s$ and $0.03\mu m^2/s$. In this work we simulate abstracted signalling and metabolic pathways both of which can involve proteins of varying size; with this in mind we choose diffusion coefficients of 1 and 0.1 representing values in-between the minimum and maximum orders of magnitude for protein diffusion. These values represent biologically realistic diffusion coefficients for the types of systems that we are interested in exploring. We reproduce a portion of the diffusion coefficients reported in [Milo and Philips, 2016] in Table 8.1.